Non-invasive Detection of Oral Cancer Using Reflectance and Fluorescence Spectroscopy

by

Sasha Alanda McGee

B.S. Chemistry
University of Maryland, Baltimore County, 2001

SUBMITTED TO THE HARVARD-MASSACHUSETTS INSTITUTE OF TECHNOLOGY DIVISION OF HEALTH SCIENCES AND TECHNOLOGY
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
IN HEALTH SCIENCES AND TECHNOLOGY

AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2008

Signature of Author: _____
Harvard-MIT Division of Health Sciences and Technology
May 27, 2008

Certified by: _____
Michael Feld
Professor of Physics
Thesis Supervisor

Accepted by: _____
Martha L. Gray, Ph.D.
Edward Hood Taplin Professor of Medical and Electrical Engineering
Director, Harvard-MIT Division of Health Sciences and Technology

NON-INVASIVE DETECTION OF ORAL CANCER USING REFLECTANCE AND FLUORESCENCE SPECTROSCOPY

by

Sasha Alanda McGee

Submitted to the Harvard-MIT Division of Health Sciences and Technology
on May 27, 2008 in Fulfillment of the Requirements for the
Doctor of Philosophy in Health Sciences and Technology

ABSTRACT

In vivo reflectance and fluorescence spectra were collected from patients with oral lesions, as well as healthy volunteers, in order to evaluate the potential of spectroscopy to serve as a non-invasive tool for the detection oral cancer. A total of 710 spectra were analyzed from 79 healthy volunteers, and 87 spectra from 67 patients. Physical models were applied to the measured spectral data in order to extract quantitative parameters relating to the structural and biochemical properties of the tissue. Data collected from healthy volunteers were used to characterize the relationship between the spectral parameters and tissue anatomy. Diagnostic algorithms for distinguishing various lesion categories were then developed using data collected from patients.

The healthy volunteer study demonstrated that tissue anatomy has a strong influence on the spectral parameters. Anatomic sites could be easily distinguished from each other despite the apparent overlap in their parameter distributions. In particular, keratinized sites (gingiva and hard palate) were significantly distinct from other anatomic sites. The results of this study provide strong evidence that a robust and accurate spectroscopic-based diagnostic algorithm for oral cancer needs to be applied in a site specific manner.

Spectral diagnostic algorithms were developed using the data collected from patients, in combination with the data collected from healthy volunteers. The diagnostic performance of the algorithms was evaluated using the area under a receiver operator characteristic curve (ROC-AUC) and the sensitivity and specificity. The diagnostic algorithms were most successful when developed and applied to data collected from a single anatomic site or spectrally similar sites, and when distinguishing visibly normal mucosa from lesions. ROC-AUC values of >0.90 could be achieved for this classification. Spectral algorithms for distinguishing benign lesions from dysplastic/malignant lesions were successfully created for the lateral surface of the tongue (ROC-AUC =0.75) and for the combination of the floor of the mouth and ventral tongue (ROC-AUC =0.71).

Thesis Supervisor: Michael Feld

Title: Professor of Physics

ACKNOWLEDGMENTS

This remarkable journey of carrying out my graduate studies at MIT could not have been possible without God, who has truly allowed me to "...run against a troop...and leap over a wall" (Psalm 18:29) to make it to this point.

First, I thank Dr. Michael Feld for his continued support and intensive involvement in my development as a scientist. I began this project literally from "scratch" in terms of my understanding of spectroscopy. He fostered a comfortable learning environment in which I could ask the most basic of questions until I finally made it to the "flat" part of the learning curve! I thank Dr. Kamran Badizadegan for his guidance and contribution to the success of this project, and Dr. Ramachandra Dasari for the laughs (and support).

While all of my colleagues in the G.R. Harrison Spectroscopy Laboratory have made this a wonderful and fun experience, no one could have contributed more than my other half, Jelena Mirkovic. Recognized as the "dynamic duo" by some, she has been my constant partner and support, scientific sounding board, and companion in adventures extending from Cambridge to Cairo. The synergy of our work experience and friendship that has evolved over these 7 years is sure to extend well beyond our time here. Additionally, I thank Dr. Chung-Chieh Yu and Condon Lau, the other half of the quantitative spectroscopy team, for their support, wonderful conversation both (scientific and non-scientific), and the laughter we shared over the course of our time together. A series of brilliant engineers contributed to the development and maintenance of the instrumentation including Luis Galindo, Jon Nazemi, Freddy Nguyen, and Timothy Brothers.

This clinical study could not have been accomplished without extensive collaboration by a hardworking team of people. The clinical work was carried out in the Department of Otolaryngology, Head and Neck Surgery at Boston Medical Center (BMC). The Principal Investigator for the project was initially Dr. Stanley Shapshay, with whom I had a wonderful

working experience. Upon his departure, Dr. Gregory Grillone took on this role and has served as such for the majority of this project. I thank him for his positive support, especially during those early years, when instrumentation problems were significant and also when great patience was required as we faced a number of research challenges. Virtually everyone in the department from administrative staff, to clinicians, specialists to residents volunteered in the Healthy Volunteer Study. I would like to acknowledge the kindness and warmth of Mike Walsh (our first volunteer for endoscopic data collection from the larynx!), Nina Leech (a wonderful presence), and Theresa Paskell (our main source of guidance in navigating the hectic clinic setting). The day-to-day hard work in linking me to the schedule of the clinicians and cases at the hospital was carried out by several research fellows. I first acknowledge Dr. Vartan Mardirossian for his outstanding work and amazing contribution to the project. He truly gave his best. I also would like to thank Dr. Alphi Elackattu, who followed him, for his hard work and dedication.

The wonderful partnerships and cooperation of several pathologists has also made this complex and logistically challenging study a success. The dedicated and hard working pathologists involved with this project included: Dr. Robert Pistey, Dr. George Gallagher, and Dr. Sadru Kabani. Dr. Pistey, a pathologist from the Department of Anatomic Pathology at BMC, served as the point person for all biopsy specimen slides. He collected, organized and did the first review of the slides prepared for each specimen. Drs. Gallagher and Kabani, oral pathologists from the Boston University Goldman School of Dental Medicine, served as additional reviewers of all the biopsy slides. During the slide review sessions, in which I had the opportunity to look through the microscope with them, they gave me great insight and understanding of the practice of oral pathology diagnosis.

I would like to acknowledge the members of my thesis committee, Dr. Roger Mark, Dr. Michael Feld, Dr. Gregory Grillone, and Dr. George Gallagher.

I thank my mom, Sacared Bodison, and dad, Elwood McGee, and my entire family for their support of my educational dreams.

TABLE OF CONTENTS:

CHAPTER 1

Overview

In this chapter we present the epidemiology of oral cancer and how it is currently detected, and then discuss how optical spectroscopy can impact this medical challenge. We then detail the specific goals of this project and conclude with a description of the organization and contents of this thesis.

1.1 Oral Cancer

1.1.1 Epidemiology of Oral Cancer

In the United States, cancers of the oral cavity and oropharynx account for approximately 3% of all malignancies in males, and 2% in females[1]. It is estimated that 35,310 new cases of oral/pharyngeal cancer will be diagnosed in 2008, and 7,590 people will die of this disease[2]. Of these cases, approximately 46% are oral cancers and 54% are pharyngeal/throat cancers[3]. Oral/pharyngeal cancer is most commonly diagnosed in persons over the age of 45 years old. The median age of diagnosis is 62 years old[4]. Men have a higher risk than women (2:1) for developing oral cancer and the incidence of oral and pharyngeal cancer tends to be higher among African Americans than among whites[5].

Tobacco and alcohol are well-established risk factors for oral cancer, and the major cause for this disease in the United States. They are estimated to account for approximately three-fourths of all oral and pharyngeal cancers in the United States[6]. In one case-control study, the relative risk in current tobacco users was found to be 6.5 as compared to controls, but with cessation gradually reduced to a comparable level after 20 years[7]. A synergistic effect between smoking and alcohol consumption results in an increased risk for oral cancer, as opposed to smoking alone[8]. Poor dentition, dietary deficiencies, and human papilloma virus (HPV) infection have also been suggested as possible causes, although there is little definitive support[9]. Considerable outdoor activity and resulting exposure to actinic radiation (sunlight), has been implicated in particular for cancer of the lip.

Early detection is critical for the successful treatment of oral/pharyngeal cancers. The 5-year survival rate falls from 82% to 52%, when the cancer is detected at a distant site, rather than when it is still confined to the primary site (localized)[10]. Unfortunately, only 33% of oral/pharyngeal cancers are localized at the time of diagnosis[10].

1.1.2 Detection of Oral Cancer

Currently the detection of oral cancer relies in large part on the vigilance of a dentist or other health care professional in identifying visible signs of disease. Alterations in surface texture (as from erosion, keratosis, induration, granularity or fissuring), color (red, white, or mixed), and pain may be associated with oral dysplasia or cancer. The presentation can be extremely variable or subtle, however, particularly in the case of early disease[8,11]. Once a suspicious lesion is identified, a small piece of tissue is cut and removed (biopsy), processed for histology, and examined under a microscope. A pathologist then renders a diagnosis based on a qualitative assessment of the tissue morphology.

This detection scheme has a number of shortcomings. First, because biopsy is an invasive procedure, only a limited number of samples can be taken to avoid excessive discomfort or stress to the patient. This means that the pathology results for a biopsy taken from one area of the lesion is assumed to be representative of the extent of the disease in the suspicious area as a whole, and often determines whether treatment is indicated[12]. Based upon his/her experience the physician selects as a site for biopsy the area most likely to show significant disease, thus there is a risk that

disease can be missed given the arbitrary nature of this process. Significant under-diagnosis has been noted to occur with incisional biopsy[12,13].

Secondly, although the entire area of the oral cavity is exposed to the carcinogens present in tobacco, diagnosis of cancer is often focused only on visibly suspicious areas. Field cancerization, the concept that multiple primary tumors develop in a field of epithelium exposed to a carcinogen, has been well described in the oral cavity[14,15]. This suggests that greater attention to the oral cavity as a whole, and not just visibly normal areas, is important for disease detection. In support of this, studies have demonstrated that visibly normal mucosa in heavy smokers or patients with lesions may also exhibit pathological changes[16,17].

Finally, the usefulness of a pathologist's diagnosis is confounded by inter- and intra-observer variability in grading, largely due to the qualitative nature of the markers used for assessment[18-20].

One alternative method for detecting oral cancer which has been explored, oral brush biopsy, suffers from a poor specificity[21]. The need for a more reliable and strategic means of diagnosing oral cancer has been highlighted in a recent review[22].

1.2 Optical Spectroscopy and Cancer Diagnosis

Greater than 85% of all cancers originate in the most superficial layer of the tissue mucosa, which is known as the epidermis. This is significant because light is able to penetrate the epidermis, and therefore optical spectroscopy may be used to

detect epithelial cancers. Studies investigating the use of optical spectroscopy for cancer diagnosis have been performed in a wide variety of tissues ranging from easily accessible tissue sites, such as the skin, oral cavity, breast, cervix, esophagus, and lungs, to more distant sites, such as the bladder, prostate, colon, stomach, and intestine[23-37].

The overall goal of these studies is to increase detection of early or localized disease, and thus the likelihood of a favorable response to treatment. The primary advantage of optical techniques is that they can non-invasively probe tissue, without the need for the biopsy or the processing of the tissue specimen required for histopathological evaluation. Research in this field is also driven by a desire to enhance the visualization of disease, to guide the clinician in the removal of malignant samples, and to reduce sampling errors in large, extensive lesions. These techniques may even reveal information beyond what is currently provided by standard detection/diagnostic methods, for example, the biochemical state of the tissue. By extracting quantitative parameters from the spectral data, diagnostic algorithms can be created, which may even permit cancer to be detected in real-time.

1.3 Project Goal and Objectives

The goal of this project is to use reflectance and fluorescence spectroscopy to non-invasively and quantitatively identify pre-cancerous or cancerous changes in the oral cavity. In order to realize this goal, we developed the following objectives:

1) To advance the instrumentation for spectral data collection
2) To develop spectral models for the accurate modeling of reflectance and fluorescence spectra from oral tissue and the reliable extraction of parameters relating to tissue morphology and biochemistry
3) To collect *in vivo* reflectance and fluorescence spectra from healthy oral tissue in healthy volunteers, and relate the extracted physical parameters to tissue anatomy
4) To collect *in vivo* reflectance and fluorescence spectra from patients with lesions and develop spectral diagnostic algorithms for identifying and categorizing oral lesions

1.4 Overview of Thesis Contents

This document consists of three main sections. The introductory section, Chapters 2-4, provides the necessary background information for understanding the architecture of the oral cavity, as well as the spectroscopic approaches and instrumentation that we will use to approach the detection of oral cancer. In the

second section, Chapters 5 and 6, we describe the collection of *in vivo* reflectance and fluorescence spectra. In Chapter 5 we describe our study of clinically normal tissue in healthy volunteers, and relate the spectroscopic parameters to the tissue anatomy. In Chapter 6 we collect spectral data from lesions in patients, and relate the spectroscopic parameters to the disease status of the tissue. In the third section, Chapters 7 and 8, we summarize the findings and conclusions of this investigation and discuss future directions. We provide a comprehensive list of abbreviations used in the subsequent chapters at the end of this thesis.

1.5 References

1. B. W. Neville and T. A. Day, "Oral cancer and precancerous lesions," *CA: A Cancer Journal for Clinicians*, **52**(4), 195-215 (2002).
2. L. A. G. Ries, D. Melbert, M. Krapcho, D. G. Stinchcomb, N. Howlader, M. J. Horner, A. Mariotto, B. A. Miller, E. J. Feuer, S. F. Altekruse, D. R. Lewis, L. Clegg, M. P. Eisner, M. Reichman and B. K. Edwards (eds), "SEER Cancer Statistics Review, 1975-2005," National Cancer Institute, 2008, accessed May 4, 2008
3. B. Rodu and P. Cole, "Oral cavity and pharynx-throat cancer in the United States, 1973-2003," *Oral Surgery Oral Medicine Oral Pathology Oral Radiology and Endodontology*, **104**(5), 653-658 (2007).
4. "SEER Cancer Stat Fact Sheets : Cancer of the Oral Cavity and Pharynx," National Cancer Institute, http://seer.cancer.gov/statfacts/html/oralcav.html, Accessed: May 3, 2008
5. L. A. G. Ries, D. Melbert, M. Krapcho, A. Mariotto, B. A. Miller, E. J. Feuer, L. Clegg, M. J. Horner, N. Howlader, M. P. Eisner, M. Reichman and B. K. Edwards (eds), "SEER Cancer Statistics Review, 1975-2004," National Cancer Institute, 2007, accessed November 12, 2007
6. W. J. Blot, J. K. McLaughlin, D. M. Winn, D. F. Austin, R. S. Greenberg, S. Prestonmartin, L. Bernstein, J. B. Schoenberg, A. Stemhagen and J. F. Fraumeni, "Smoking And Drinking In Relation To Oral And Pharyngeal Cancer," **48**(11), 3282-3287 (1988).
7. F. Lewin, S. E. Norell, H. Johansson, P. Gustavsson, J. Wennerberg, A. Biorklund and L. E. Rutqvist, "Smoking tobacco, oral snuff, and alcohol in the etiology of squamous cell carcinoma of the head and neck - A population-based case-referent study in Sweden," *Cancer*, **82**(7), 1367-1375 (1998).
8. J. J. Sciubba, "Oral cancer and its detection - History-taking and the diagnostic phase of management," **132**(12S-18S (2001).
9. L. Dobrossy, "Epidemiology of head and neck cancer: Magnitude of the problem," *Cancer and Metastasis Reviews*, **24**(1), 9-17 (2005).
10. A. Jemal, R. Siegel, E. Ward, Y. Hao, J. Xu, T. Murray and M. J. Thun, "Cancer Statistics, 2008," *CA: A Cancer Journal for Clinicians*, **58**(2), 71-96 (2008).
11. E. E. Vokes, R. R. Weichselbaum, S. M. Lippman and W. K. Hong, "Medical Progress - Head and Neck-Cancer," *New England Journal of Medicine*, **328**(3), 184-194 (1993).
12. M. Pentenero, M. Carrozzo, M. Pagano, D. Galliano, R. Broccoletti, C. Scully and S. Gandolfo, "Oral mucosal dysplastic lesions and early squamous cell carcinomas: underdiagnosis from incisional biopsy," *Oral Diseases*, **9**(2), 68-72 (2003).
13. J. J. Lee, H. C. Hung, S. J. Cheng, C. P. Chiang, B. Y. Liu, C. H. Yu, J. H. Jeng, H. H. Chang and S. H. Kok, "Factors associated with underdiagnosis from incisional biopsy of oral leukoplakic lesions," *Oral Surgery Oral*

Medicine Oral Pathology Oral Radiology and Endodontology, **104**(2), 217-225 (2007).

14. D. P. Slaughter, H. W. Southwick and W. Smejkal, "Field cancerization in oral stratified squamous epithelium; clinical implications of multicentric origin," *Cancer*, **6**(5), 963-8 (1953).
15. F. Cianfriglia, D. A. Di Gregorio and A. Manieri, "Multiple primary tumours in patients with oral squamous cell carcinoma," **35**(2), 157-163 (1999).
16. P. J. Thomson, "Field change and oral cancer: new evidence for widespread carcinogenesis?," *International Journal of Oral and Maxillofacial Surgery*, **31**(3), 262-266 (2002).
17. N. Ayan, I. Ayan, C. Alatli, S. D. Guler, C. Dalkilic, F. Cinar and O. Dogan, "P53 overexpression in normal oral mucosa of heavy smokers," *Journal of Experimental & Clinical Cancer Research*, **19**(4), 525-529 (2000).
18. D. J. Fischer, J. B. Epstein, T. H. Morton and S. M. Schwartz, "Reliability of histologic diagnosis of clinically normal intraoral tissue adjacent to clinically suspicious lesions in former upper aerodigestive tract cancer patients," *Oral Oncology*, **41**(5), 489-496 (2005).
19. D. J. Fischer, J. B. Epstein, T. H. Morton and S. M. Schwartz, "Interobserver reliability in the histopathologic diagnosis of oral pre-malignant and malignant lesions," *Journal of Oral Pathology & Medicine*, **33**(2), 65-70 (2004).
20. I. vanderWaal, K. P. Schepman, E. H. vanderMeij and L. E. Smeele, "Oral leukoplakia: a clinicopathological review," *Oral Oncology*, **33**(5), 291-301 (1997).
21. T. W. J. Poate, J. A. G. Buchanan, T. A. Hodgson, P. M. Speight, A. W. Barrett, D. R. Moles, C. Scully and S. R. Porter, "An audit of the efficacy of the oral brush biopsy technique in a specialist Oral Medicine unit," *Oral Oncology*, **40**(8), 829-834 (2004).
22. R. Tagg, M. Asadi-Zeydabadi and A. D. Meyers, "Biophotonic and other physical methods for characterizing oral mucosa," *Otolaryngologic Clinics of North America*, **38**(2), 215-240 (2005).
23. S. C. Park, S. J. Lee, H. Namkung, H. Chung, S. H. Han, M. Y. Yoon, J. J. Park, J. H. Lee, C. H. Oh and Y. A. Woo, "Feasibility study for diagnosis of stomach adenoma and cancer using IR spectroscopy," *Vibrational Spectroscopy*, **44**(2), 279-285 (2007).
24. S. K. Chang, Y. N. Mirabal, E. N. Atkinson, D. Cox, A. Malpica, M. Follen and R. Richards-Kortum, "Combined reflectance and fluorescence spectroscopy for in vivo detection of cervical pre-cancer," *Journal of Biomedical Optics*, **10**(2), - (2005).
25. C. M. Krishna, N. B. Prathima, R. Malini, B. M. Vadhiraja, R. A. Bhatt, D. J. Fernandes, P. Kushtagi, M. S. Vidyasagar and V. B. Kartha, "Raman spectroscopy studies for diagnosis of cancers in human uterine cervix," **41**(1), 136-141 (2006).
26. T. M. Breslin, F. S. Xu, G. M. Palmer, C. F. Zhu, K. W. Gilchrist and N. Ramanujam, "Autofluorescence and diffuse reflectance properties of

malignant and benign breast tissues," *Annals of Surgical Oncology*, **11**(1), 65-70 (2004).

27. A. L. Carlson, A. M. Gillenwater, M. D. Williams, A. K. El-Naggar and R. R. Richards-Kortum, "Confocal microscopy and molecular-specific optical contrast agents for the detection of oral neoplasia," *Technology in Cancer Research & Treatment*, **6**(5), 361-374 (2007).
28. M. P. L. Bard, A. Amelink, M. Skurichina, V. N. Hegt, R. P. W. Duin, H. J. C. M. Sterenborg, H. C. Hoogsteden and J. G. J. V. Aerts, "Optical spectroscopy for the classification of malignant lesions of the bronchial tree," *Chest*, **129**(4), 995-1001 (2006).
29. P. Crow, A. Molckovsky, N. Stone, J. Uff, B. Wilson and L. M. Wongkeesong, "Assessment of fiberoptic near-infrared Raman spectroscopy for diagnosis of bladder and prostate cancer," *Urology*, **65**(6), 1126-1130 (2005).
30. A. S. Haka, K. E. Shafer-Peltier, M. Fitzmaurice, J. Crowe, R. R. Dasari and M. S. Feld, "Diagnosing breast cancer by using Raman spectroscopy," *Proceedings of the National Academy of Sciences of the United States of America*, **102**(35), 12371-12376 (2005).
31. D. C. Malins, N. L. Polissar, K. Nishikida, E. H. Holmes, H. S. Gardner and S. J. Gunselman, "The Etiology and Prediction of Breast-Cancer - Fourier Transform-Infrared Spectroscopy Reveals Progressive Alterations in Breast DNA Leading to a Cancer-Like Phenotype in a High Proportion of Normal Women," *Cancer*, **75**(2), 503-517 (1995).
32. J. H. Ali, W. B. Wang, M. Zevallos and R. R. Alfano, "Near infrared spectroscopy and imaging to probe differences in water content in normal and cancer human prostate tissues," *Technology in Cancer Research & Treatment*, **3**(5), 491-497 (2004).
33. G. Bourg-Heckly, J. Blais, J. J. Padilla, O. Bourdon, J. Etienne, F. Guillemin and L. Lafay, "Endoscopic ultraviolet-induced autofluorescence spectroscopy of the esophagus: Tissue characterization and potential for early cancer diagnosis," *Endoscopy*, **32**(10), 756-765 (2000).
34. L. D. Nunes, A. A. Martin, L. Silveira and M. Zampieri, "FT-Raman spectroscopy study for skin cancer diagnosis," *Spectroscopy-an International Journal*, **17**(2-3), 597-602 (2003).
35. L. Brancaleon, A. J. Durkin, J. H. Tu, G. Menaker, J. D. Fallon and N. Kollias, "In vivo fluorescence spectroscopy of nonmelanoma skin cancer," *Photochemistry and Photobiology*, **73**(2), 178-183 (2001).
36. A. Molckovsky, L. M. W. K. Song, M. G. Shim, N. E. Marcon and B. C. Wilson, "Diagnostic potential of near-infrared Raman spectroscopy in the colon: differentiating adenomatous from hyperplastic polyps," *Gastrointestinal Endoscopy*, **57**(3), 396-402 (2003).
37. G. Zonios, L. T. Perelman, V. M. Backman, R. Manoharan, M. Fitzmaurice, J. Van Dam and M. S. Feld, "Diffuse reflectance spectroscopy of human adenomatous colon polyps in vivo," *Applied Optics*, **38**(31), 6628-6637 (1999).

CHAPTER 2

Introduction to the Oral Cavity

We begin this chapter with an overview of the normal macroscopic and microscopic architecture of the oral cavity. Next, we provide an overview of the process of cancer development and describe commonly encountered lesions. Finally, we end with a description of the primary treatment for oral cancer.

2.1 Anatomy of the Oral Cavity

2.1.1 Macroscopic Anatomy

The oral cavity includes the lips, anterior two-thirds of the tongue, gingiva (gums), buccal mucosa (lining inside the cheeks and lips), floor of the mouth (area under the tip of the tongue in the anterior region of the mouth), hard palate, and retromolar trigone (area of the gums distal to the wisdom teeth). The oropharynx includes the soft palate, the base of the tongue, and the tonsils. Figure 2.1 depicts the anatomic sites of the oral cavity and oropharynx[1].

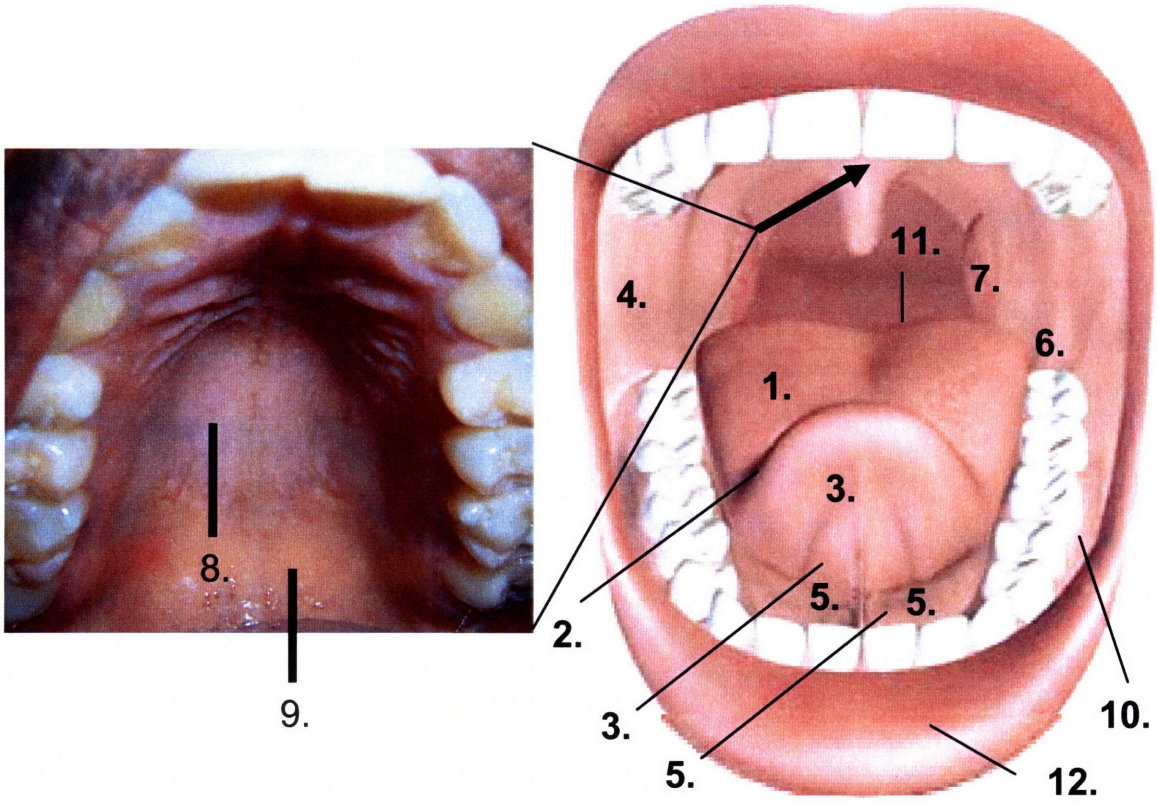

Figure 2.1 Diagram and photograph depicting the various anatomic sites of the oral cavity and oropharynx. 1. Dorsal surface of the tongue, 2. Lateral surface of the tongue, 3. Ventral surface of the tongue, 4. Buccal mucosa, 5. Floor of the mouth, 6. Retromolar trigone, 7. Tonsils, 8. Hard palate, 9. Soft palate, 10. Gingiva, 11. Base of tongue, 12. Lips. Diagram modified from http://www.3dscience.com/3D_Images/Human_Anatomy/Digestive/Open_Mouth.php

Unless otherwise noted, as in the diagram used in Figure 2.1, all photographs that are shown in this chapter are derived from samples collected during the course of our study of the oral cavity.

2.1.2 Microscopic Anatomy

The mucosa of the oral cavity consists of two layers: an outer (most superficial) layer of stratified squamous epithelium, and an underlying layer of connective tissue known as the lamina propria. The two layers are separated by a structure known as the basement membrane or basal lamina. Figure 2.2 shows a hemotoxylin and eosin (H&E) stained specimen taken from the hard palate, showing the normal tissue micro-architecture.

The mucosa of the oral cavity varies for different anatomic sites and is generally divided into three categories based on specific structural features which are related to the function of the particular site[2]. Lining mucosa covers the lips, buccal mucosa (BM), soft palate (SP), ventral surface of the tongue (VT), and floor of the mouth (FM). Lining mucosa is non-keratinized, flexible, and easily stretched. Masticatory mucosa is keratinized [Fig. 2.2] and covers those sites which are exposed to the intense forces of mastication: the hard palate (HP) and gingiva (GI). The dorsal surface (anterior two-thirds) of the tongue (DT) is characterized by specialized mucosa. Specialized mucosa consists of a mixture of keratin and surface projections known as papillae[3].

Epithelium

The epithelial layer protects the tissue from microbial invasion and mechanical damage. The most inferior portion consists of a layer of actively dividing cells, 1-2 cell layers in thickness, which are located just superficial to the basement membrane. This zone of epithelial cells is known as the basal cell layer. Basal cells are tightly packed and undifferentiated, and eventually progress upwards to form the

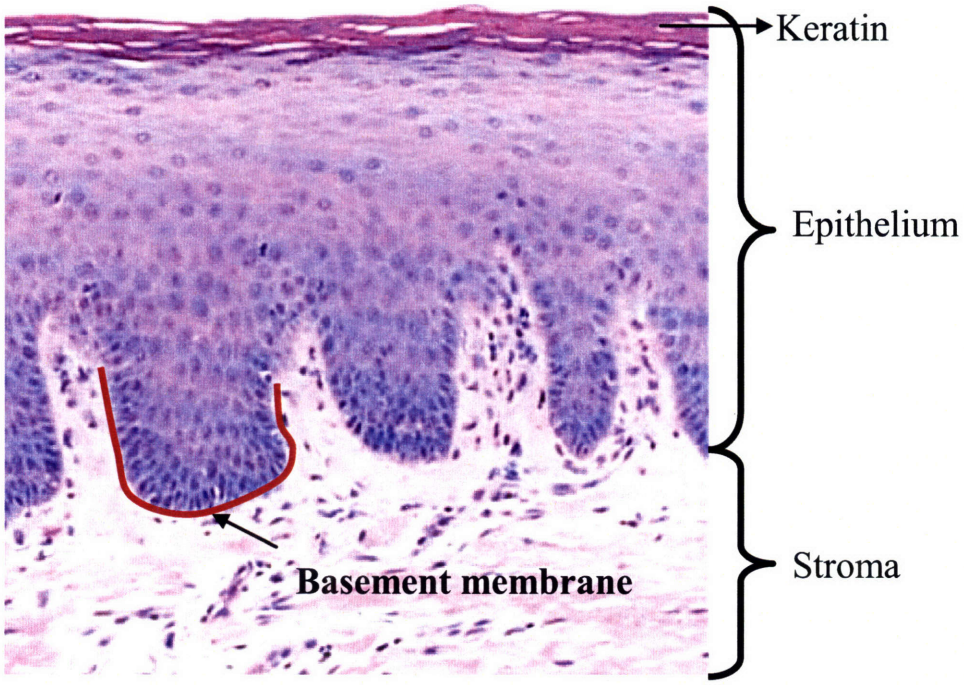

Figure 2.2 A H&E stained tissue specimen from the hard palate depicting the epithelium with overlying keratin, the basement membrane, and the lamina propria.

most superficial layer. The cells progressively become larger and more differentiated as they advance towards the tissue surface. At the tissue surface, the epithelial cells are dead, and are shed during activities such as chewing and talking. These cells are replaced by the continuous proliferation and upward migration of the basal cell layer[4].

In keratinized epithelium (GI and HP), the basal cells produce keratin intermediate filaments. These filaments become densely packed and the nuclear-to-cytoplasmic ratio decreases as the cells migrate upwards, such that eventually the cells have eliminated all their organelles. At the surface, the cells are no longer alive and the keratin is released into the intracellular spaces[2]. If the nucleus is retained, this is referred to as parakeratosis, while if absent, the pattern of keratinization is referred to as orthokeratosis. Non-keratinized epithelium is usually thicker than keratinized epithelium. Both keratinized and non-keratinized epithelia contain keratin. Distinction between these two types of epithelia is based on differences in the types of keratins present[2].

The specialized mucosa of the DT possesses a number of unique features, related to its important functions. The base of the tongue (posterior two-thirds of the tongue) is often distinguished from the anterior two-thirds because of their unique embryologic origin and morphology. In this thesis, we will use DT to refer to the anterior two-thirds, and explicitly refer to the posterior portion as the base of the tongue (BT). The DT contains four types of papillae, small protuberances, (some of which contain taste buds) in a distinct spatial pattern. The fungiform, circumvallate, and foliate papillae are non-keratinized, however the filiform papillae are covered by keratinized epithelium[4].

Basement Membrane

The basement membrane serves as a scaffold for regeneration of the epithelium, influences cell proliferation and growth, and controls cell polarity[5]. The

basement membrane is composed of collagen type IV and VII, glycoproteins, and a number of other types of fibrils[2]. The basement membrane is not a flat structure, but rather undulates because the interface between the epithelium and lamina propria is corrugated. This interface is characterized by inderdigitating epithelial ridges and connective tissue papillae.

Lamina Propria and Submucosa

The lamina propria contains blood vessels, fat, nerves and glands, but the bulk of this compartment is composed of collagen (mostly type I and III) produced by fibroblasts, the most abundant cell type found in this layer. Also found within this layer are elastin, white blood cells, macrophages, and lymphatics[3]. The lamina propria provides support for the overlying epithelium.

Deep to the lamina propria is the submucosa, which may be composed of fat (SP, BM, and labial mucosa), muscle (lips, BM, tongue, and SP), bone (GI and anterior portion of the HP), or glands (lips, BM, HP, SP, and tongue)[3]. The GI and HP are unique in that neither contains a submucosa; but rather, the fibers of the lamina propria are attached directly to bone[4].

2.2 Oral Cavity Pathology

2.2.1 Carcinogenesis

A neoplasm ("new growth") represents a virtually autonomous mass of tissue characterized by uncontrolled, uncoordinated growth, which is independent of normal

growth stimuli and plays no role in the normal tissue function[5]. Carcinogenesis is the multi-step process by which the original transformed cell expands and progressively transforms into the malignant phenotype. Malignancy is characterized by features such as accelerated growth, invasiveness, and the ability to form metastases[5]. Based on a microscopic assessment of a biopsied tissue specimen, oral lesions are commonly classified as benign, dysplastic, or malignant (cancer).

Benign lesions are characterized by changes which are abnormal, but are considered less likely to progress to the malignant phenotype. Common benign changes associated with oral lesions include hyperkeratosis (increased thickness of the overlying keratin) and hyperplasia (proliferation) of keratinocytes (known as acanthosis).

For oral lesions, epithelial dysplasia has been defined as "a pre-cancerous lesion of stratified squamous epithelium characterized by cellular atypia and loss of normal maturation..."[6] Table 2.1 summarizes the abnormal histological features associated with dysplasia[6]. The presence of dysplasia is considered an indication of increased risk of malignant transformation as compared to non-dysplastic epithelium. Not all dysplastic lesions will eventually become malignant, however, and some lesions may even regress[7]. Dysplastic lesions are distinguished from frank cancer by their confinement to a region superficial to the basement membrane. Dysplastic lesions are sometimes subdivided into three categories (mild, moderate and severe), based on whether the basal cell layer and other changes extend up to one third, two-thirds, or greater than two-thirds, respectively, of the thickness of the epithelium[8].

Histological Criteria for Diagnosing Epithelial Dysplasia
• Nuclear hyperchromatism (increased nuclear staining)
• Cellular and nuclear pleomorphism (variation in nuclear shape and size)
• Enlarged nuclei
• Irregular epithelial stratification
• Increased number of mitotic figures
• Mitotic figures that are abnormal in form
• The presence of mitotic figures in the superficial half of the epithelium
• Loss of polarity of basal cells
• Increased nuclear-cytoplasmic ratio
• Drop-shaped rete ridges
• Loss of intercellular adherence
• The presence of more than one layer of cells having a basaloid appearance
• Keratinization of single cells or cell groups in the prickle cell layer

Table 2.1 A description of the major histological features associated with dysplasia.

Epithelial cancer (also known as squamous cell carcinoma (SCC)) develops once the neoplastic cells penetrate the basement membrane and invade the underlying stroma. Approximately 92% of all oral cancers are SCCs[9]. A term used to describe an intermediate stage in which there is uncontrolled growth of epithelial cells throughout the full thickness of the epithelium, but without stromal invasion, is known as carcinoma *in situ* (CIS)[6].

Figure 2.3 shows a series of microscopic specimens obtained from the floor of the mouth showing the histological changes with progression from mild dysplasia to cancer.

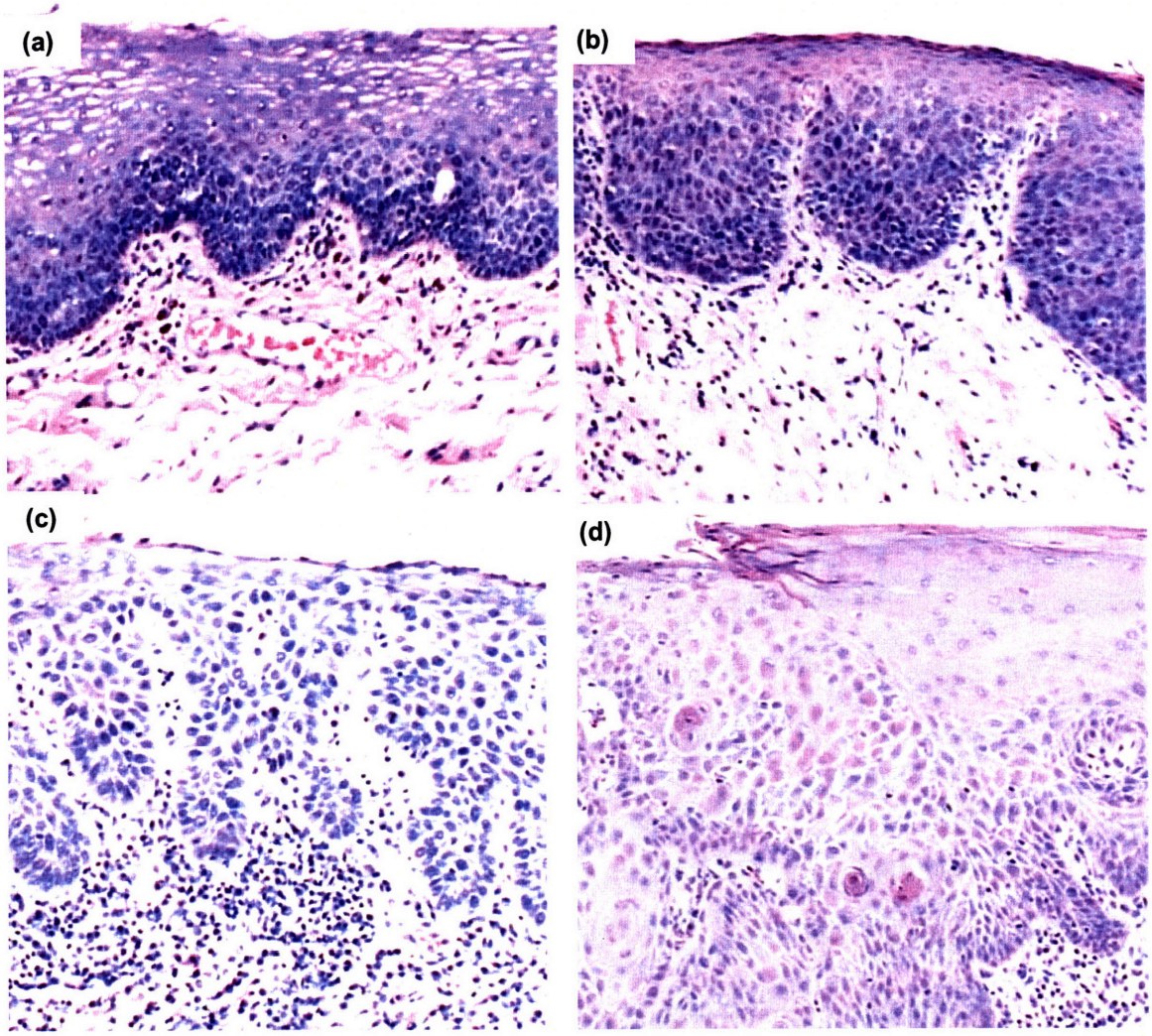

Figure 2.3 H&E stained microscopic specimens from the floor of the mouth showing (a) mild dysplasia, (b) moderate dysplasia, (c) severe dysplasia, and (d) cancer. The progression of dysplasia is characterized by proliferation and expansion of the actively dividing cells of the basal cell layer to more superficial regions of the epithelium. Finally, with cancer, the rapidly proliferating cells breech the basement membrane (note the loss of this structure) and invade the underlying stroma, where they continue to proliferate.

Figure 2.4 shows photographs of 2 different regions of the epithelium of a dysplastic lateral tongue lesion demonstrating the nuclear and cellular pleomorphism [Fig. 2.4(a)] and presence of mitotic figures in the superficial region of the epithelium [Fig. 2.4(b)].

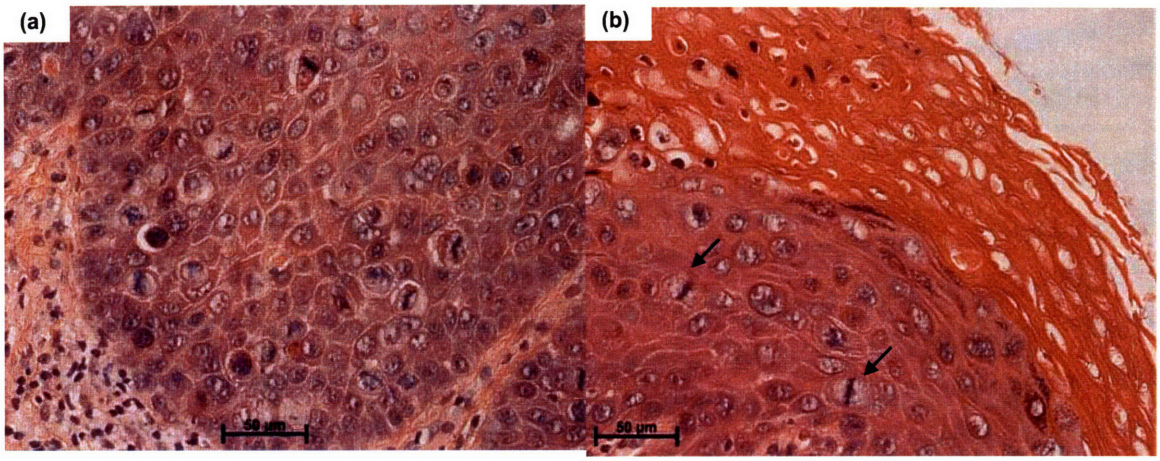

Figure 2.4 Two photographs of a H&E stained microscopic specimen showing a dysplastic lesion from the lateral tongue. (a) The basal cells and the epithelial-stromal junction. The variation in the size and shape of the nuclei and cells (pleomorphism) can be appreciated. (b) Superficial region of the epithelium. The presence of actively dividing cells near the tissue surface is evidenced by the presence of mitotic figures (2 are indicated by the arrows). This represents an abnormal maturation pattern.

The tongue is the most common location of cancer development in the oral cavity and oropharynx, followed by the floor of the mouth[10]. In a study of a series of 200 patients with oral SCC, 26% had a primary tumor on the lateral border of the tongue, 17% in the ventral tongue/floor of mouth region, and 15% solely in the floor of mouth[11]. The most common sites of metastasis of oral cancers include the mediastinal lymph nodes, lungs, liver, and bones[5]. Invasive lesions can be described as poorly, moderately or well-differentiated based on the degree of keratinization (poor indicating the least amount of keratinization). In addition to histological grading, the TNM system has been used for staging oral lesions. Using this classification, the tumor size (T), status of the lymph nodes with respect to metastasis (N), and the presence or absence of distant metastases (M) are evaluated[12].

2.2.2 *Lesions of the Oral Cavity*

There are two lesions commonly encountered in the oral cavity: leukoplakia ("white plaque") and erythroplakia ("red plaque").

Leukoplakia

Leukoplakia, the most commonly observed suspicious lesion found in the oral cavity, is a clinical term which refers to a predominantly white lesion of the oral mucosa that cannot be wiped off or attributed to any specific disease etiology[13]. Non-idiopathic causes of white lesions include candidiasis (fungal infection by *Candida albicans*) and lichen planus (an inflammatory disease often producing lesions on the oral mucosa)[14]. Under microscopic examination, white lesions can range in appearance from benign areas of hyperkeratosis, to dysplasia, or frank cancer. The white appearance of the plaque is due to the keratin filaments, which scatter all wavelengths of light. Upon microscopic evaluation, most cases of leukoplakia show only benign changes. In a study of 3,256 biopsied leukoplakias, 80.2% showed benign changes, particularly hyperkeratosis and hyperplasia. Of the remaining cases, 16.7% showed dysplasia/CIS and 3.1% were diagnosed as infiltrating SCC[15]. Interestingly, the frequency of disease findings for white lesions was found to vary among different sites in the mouth. The sites with the highest incidence included the floor of the mouth (42.9%), tongue (24.2%), lip (24%), and palate (18.8%).

Reports vary widely on the rate of transformation of benign leukoplakia to SCC. In a study of 257 patients with leukoplakia studied for an average of 7.1 years,

45 patients (17.5%) with an initial diagnosis of benign hyperkeratosis subsequently developed squamous cell carcinomas in an average time of 8.1 years[16]. The transformation rate for leukoplakia, however, varies widely across studies, ranging from 0.9% to 17.5%[7].

Erythroplakia

Erythroplakia is a less commonly observed lesion, and presents as a red, velvety, flat or slightly depressed lesion. This lesion has a prevalence between 0.02% and 0.83%[17]. Like leukoplakia, this term is simply a clinical description of the lesion. The description is replaced once the diagnosis is rendered. The diagnosis at the time of initial biopsy of erythroplakia tends to be far more severe than that for leukoplakia. In a study by Shafer and Waldron of 65 cases of erythroplakia, all of the cases showed the presence of dysplasia (49%) or invasive SCC (51%)[18]. Under the microscope these red lesions frequently display a superficially eroded epithelium, an intense inflammatory reaction within the stroma with vascular dilation and signs of early or late cancerous changes[5]. The red color of these lesions is due to increased vascularity and inflammation. Erythroplakia is commonly found on the soft palate and floor of the mouth[17]. Some red lesions are mixed with white areas of keratosis and are referred to as erythroleukoplakias.

2.3 Treatment for Oral Cancer

Surgical excision is the primary means of treating oral cancer[19]. Overall disease-free survival at 5 years is 58%[20]. Successful treatment has been found to be significantly affected by the histopathological grade at the time of diagnosis and the stage of disease (according to the TNM classification system)[20]. There is a relatively high recurrence rate for oral lesions and surgical excision does not reduce this risk[19]. In addition, the development of second (metachronous) primary tumors is common, particularly within the first twelve months of treatment[21].

2.4 References

1. Zygote Media Group, Inc.,"Open Mouth", http://www.3dscience.com/3D_Images/Human_Anatomy/Digestive/Open_Mouth.php, accesesed December 2007
2. A. R. Ten Cate, *Oral histology : development, structure, and function*, Mosby, St. Louis, 1998
3. T. A. Winning and G. C. Townsend, "Oral mucosal embryology and histology," *Clinics in Dermatology*, **18**(5), 499-511 (2000).
4. S. S. Sternberg, *Histology for pathologists*, Lippincott Williams & Wilkins, New York, 1997
5. R. S. Cotran, V. Kumar, T. Collins and S. L. Robbins, *Robbins pathologic basis of disease*, Saunders, Philadelphia, 1999
6. J. Pindborg, P. Reichart, C. Smith and I. van der Waal, *World Health Organization: histological typing of cancer and precancer of the oral mucosa*, Springer-Verlag, Berlin, 1997
7. J. Reibel, "Prognosis of oral pre-malignant lesions: Significance of clinical, histopathological, and molecular biological characteristics," *Critical Reviews in Oral Biology and Medicine*, **14**(1), 47-62 (2003).
8. S. Warnakulasuriya, "Histological grading of oral epithelial dysplasia: revisited," *Journal of Pathology*, **194**(294-297 (2001).
9. J. P. Shah, N. W. Johnson and J. G. Batsakis, *Oral cancer*, Thieme Medical Publishers, Inc., New York, NY, 2003
10. B. W. Neville and T. A. Day, "Oral cancer and precancerous lesions," *CA: A Cancer Journal for Clinicians*, **52**(4), 195-215 (2002).
11. J. A. Woolgar, S. Rogers, C. R. West, R. D. Errington, J. S. Brown and E. D. Vaughan, "Survival and patterns of recurrence in 200 oral cancer patients treated by radical surgery and neck dissection," **35**(3), 257-265 (1999).
12. *AJCC (American Joint Committee on Cancer) Manual for Staging of Cancer*, F. L. Greene, A. G. Fritz, C. M. Balch, D. G. Haller, D. L. Page, I. D. Fleming and M. Morrow (Eds). Springer-Verlag, New York, 2002
13. R. A. Cawson and E. W. Odell, *Essentials of oral pathology and oral medicine*, Churchill Livingstone, New York, 2002
14. I. vanderWaal, K. P. Schepman, E. H. vanderMeij and L. E. Smeele, "Oral leukoplakia: a clinicopathological review," *Oral Oncology*, **33**(5), 291-301 (1997).
15. C. A. Waldron and W. G. Shafer, "Leukoplakia Revisited - Clinicopathologic Study 3256 Oral Leukoplakias," *Cancer*, **36**(4), 1386-1392 (1975).
16. S. Silverman, M. Gorsky and F. Lozada, "Oral Leukoplakia And Malignant Transformation - A Follow-Up-Study Of 257 Patients," **53**(3), 563-568 (1984).
17. P. A. Reichart and H. P. Philipsen, "Oral erythroplakia - a review," **41**(6), 551-561 (2005).
18. W. G. Shafer and C. A. Waldron, "Erythroplakia Of Oral Cavity," *Cancer*, **36**(3), 1021-1028 (1975).

19. G. Lodi and S. Porter, "Management of potentially malignant disorders: evidence and critique," *Journal of Oral Pathology & Medicine*, **37**(2), 63-69 (2007).
20. D. Kademani, R. B. Bell, S. Bagheri and E. Holmgren, "Prognostic factors in intraoral squamous cell carcinoma: The influence of histologic grade," *Journal of Oral and Maxillofacial Surgery*, **63**(11), 1599-1605 (2005).
21. J. A. Woolgar, S. Rogers, C. R. West, R. D. Errington, J. S. Brown and E. D. Vaughan, "Survival and patterns of recurrence in 200 oral cancer patients treated by radical surgery and neck dissection," *Oral Oncology*, **35**(3), 257-265 (1999).

CHAPTER 3

Introduction to Spectroscopy: Methods and Models

We begin this chapter by describing the advantages and motivation for using optical techniques to detect malignancy. Following this, we introduce the major optical spectroscopy, imaging, and microscopy modalities that have been applied to the oral cavity. For each technique, we briefly summarize the relevant literature on the application of the technique to the evaluation of normal or diseased oral tissue. The results of other investigators will be described in further detail and compared with our own findings in the ensuing chapters. Finally, we introduce the two physical models we use to analyze the spectral data and extract physical parameters related to tissue morphology and biochemistry: Diffuse Reflectance Spectroscopy (DRS) and Intrinsic Fluorescence Spectroscopy (IFS). Light Scattering Spectroscopy (LSS), a technique previously used by our laboratory will also be discussed.

3.1 Spectroscopy and Cancer Diagnosis: Motivation

A number of structural and metabolic changes are associated with the development of dysplasia and cancer that significantly alter the normal architecture and biochemical properties of tissue[1-7]. Many of the same tissue components affected by the disease process also interact with light through processes such as scattering, absorption, and fluorescence. Therefore, as these components are altered, we expect the impact of disease to produce changes in the light emitted from the tissue.

There is an extensive body of literature describing spectroscopic changes associated with disease (i.e. dysplasia, cancer) in epithelial tissue[8-14]. Table 3.1 provides an overview of the findings associated with benign, dysplastic, and malignant lesions, and how these changes can produce differences in the spectroscopic properties of the tissue. The actual spectroscopic changes observed with carcinogenesis will represent a combination of the individual (or dominant) changes listed, and depend on factors such as the tissue under study, and the methods and instrumentation used to probe the tissue. Frequent findings reported in the literature with regard to the spectral changes associated with disease include increased hemoglobin absorption and nicotinamide adenine dinucleotide in the reduced form (NADH) fluorescence, and a decrease in the magnitude of the scattering and collagen fluorescence[9,13,15].

Changes Associated with Benign/Dysplastic/Malignant Lesions	Spectroscopic Signature
Hyperplasia, nuclear crowding	Increased scattering by small particles
Increased nuclear/cytoplasm ratio	Increased scattering by small particles
Degradation of extracellular matrix (collagen)	Decreased scattering from collagen Decreased fluorescence from collagen
Loss of basement membrane (collagen)	Decreased scattering from collagen Decreased fluorescence from collagen
Angiogenesis	Increased hemoglobin absorption
Decreased hemoglobin oxygenation (hypoxia)	Oxyhemoglobin absorption peaks decrease, deoxyhemoglobin absorption peaks increase, (increased tissue hemoglobin oxygenation)
Increased metabolic activity	Increased NADH fluorescence Decreased FAD fluorescence
Inflammation	Increased hemoglobin absorption
Hyperkeratosis	Increased scattering

Table 3.1 The morphological and biochemical changes associated with benign, dysplastic, and malignant lesions and their spectroscopic correlate. NADH: reduced nicotinamide adenine dinucleotide, FAD: flavin adenine dinucleotide.

There are a variety of methods to detect these optical contrasts, and the major techniques and approaches will be described in the subsequent section.

3.2 Optical Techniques and their Application to the Oral Cavity

Upon irradiating tissue with light, several types of optical signals are generated from interactions within the tissue; however, three important types are scattering, absorption, and fluorescence. These processes form the basis for the spectroscopic and imaging techniques used to non-invasively investigate oral tissue. In the following subsections we review the major optical techniques and give a brief overview of their application in the oral cavity specifically. We focus primarily on

studies that have been done *in vivo* in humans, as they are most relevant to the work described in this thesis.

3.2.1 Diffuse Reflectance Spectroscopy: Elastic Scattering and Absorption

Diffuse reflectance spectroscopy (DRS) examines the changes in the properties of light due to the interplay of elastic scattering and absorption within a medium. This technique provides information about the bulk architecture of a small volume of tissue. Broad bandwidth light is delivered to the tissue and the re-emitted light is collected (both usually with an optical fiber probe). The returned light is then dispersed by a spectrograph and detected by a CCD. The simple, inexpensive instrumentation is one of the major advantages of diffuse reflectance spectroscopy. In addition, the measurements can be performed very rapidly.

Elastic Scattering

Elastic scattering occurs when the path of light is redirected as a result of variations in the refractive indices of cellular and extracellular constituents within the tissue, with respect to their surrounding medium. The refractive index is defined as the ratio of the velocity of light as it travels through a vacuum relative to the velocity as it travels through a medium. In elastic scattering, the scattered light ray has the same energy as the incident ray. Scattering particles present within tissue include nuclei, mitochondria, collagen, elastin, and keratin[16]. The extent to which light is scattered is a function of the density of particles, the sizes of the particles relative to the wavelength of the light, and the ratio of the refractive indices of the particles

relative to that of the surrounding medium. The pattern of the scattered light largely reflects those structures that have the highest refractive index mismatch compared to their surrounding. However, because of the complex mixture of components present within tissue (and thus inhomogeneities in refractive index), many structures ultimately contribute to the scattering.

Light that enters tissue can be re-emitted after undergoing only a single scattering event or after multiple scattering events. In principle, evaluation of singly-scattered light, Light Scattering Spectroscopy (LSS), can provide in depth information about specific structures localized in the most superficial layers of a tissue (Section 3.3.3). Diffusely scattered light refers to elastic scattering in which the incident light undergoes multiple scattering events to such an extent that it becomes randomized in direction. Because the measured signal represents light that has sampled a variety of paths and depths within the tissue, the diffuse reflectance spectrum represents an average measure of the properties over a volume of tissue. Specular reflectance can occur if the incident light rays encounter a microscopically smooth surface (i.e. a mirror). In this case, the light rays are not reflected in a broad range of directions as in diffuse reflectance.

Absorption

Molecules and atoms can absorb incident light at specific energies, converting the light's energy into internal energy and ultimately, heat. The structure of the molecule determines which incident frequencies will be absorbed. When tissue is irradiated, the process of absorption causes a reduction in the amount of the incident

light which returns to the surface to be collected. Table 3.2 lists the major tissue absorbers in the ultraviolet (UV) and visible regions of the spectrum[16-19].

Absorber	Absorption Peak(s) / Range [nm]
Nucleic Acids: DNA ,RNA	258
Amino acid: Tyrosine	275
Amino acid: Tryptophan	280
Amino acid: Phenylalanine	260
Oxyhemoglobin	415, 542,576
Deoxyhemoglobin	433, 556
β-Carotene	<300nm, ~450
Melanin	400-700

Table 3.2 Summary of the major absorbers present in tissue and their absorption maxima or absorption range.

In the UV region of the spectrum, nucleic acids and amino acids are the major absorbers. Of the amino acids, tryptophan exhibits the most intense absorption. Light is readily absorbed in the UV region and thus is able to penetrate only one or two cell layers deep into the tissue. Hemoglobin is the principal absorber in tissue in the visible region of the spectrum and strongly absorbs blue light. The removal of the blue component from the beam of incident white light results in its characteristic red color. Melanin and β-carotene are additional absorbers in the visible region. Visible light typically penetrates tissue to a depth of 0.5-2.5 mm[20]. In the wavelength region from 600-1500 nm, known as the 'diagnostic window', scattering prevails over absorption, and light can penetrate as deep as 8-10 mm before being collected[20].

Diffuse Reflectance Studies in the Oral Cavity

There have been relatively few reports describing the use of diffuse reflectance spectroscopy *in vivo* in humans for the detection of oral malignancy [9,21-23].

In one study, a preliminary examination of this technique was presented and the data included only 6 people[22]. In another, the researchers compared *ex vivo* tissue data from patients with *in vivo* data from healthy volunteers[23]. Sharwani *et al.* conducted an *in vivo* study in which they collected 25 white light spectra from 25 patients with leukoplakia and analyzed the data from 340-800 nm[21]. Their system used an optical fiber probe to convey light from a xenon-arc lamp to the tissue and also the emitted light to the CCD-based detection system. The data included 4 normal sites, 10 benign sites, 10 dysplastic sites, and 1 carcinoma *in-situ* (CIS). Using linear discriminant analysis followed by leave-one-out cross-validation, they obtained a sensitivity and specificity of 72.7% and 75%, respectively, for separating normal from dysplastic sites. Recently, Amelink *et al.* compared clinically normal versus cancerous sites in 31 patients using an optical fiber probe designed to capture the returned signal specifically from a superficial layer of the tissue (~350 μm)[9]. Using a model-based approach to extract physical parameters, they found that the cancerous sites demonstrated decreased scattering and oxygen saturation values, and increased blood content and scattering slope. A study by de Veld *et al.* examined diffuse reflectance spectra from oral tissue in combination with fluorescence from 172 oral lesions and 70 healthy volunteers[24]. They analyzed the spectra from 400-700 nm using principal component analysis (PCA), after applying various normalization methods. Multiple classification algorithms were also tested for distinguishing various subsets of the data. They found that lesions could successfully be distinguished from healthy tissue, and the best separation for distinguishing benign lesions from dysplastic/malignant lesions yielded a sensitivity and specificity of 69% and 77%, respectively.

3.2.2 *Raman Spectroscopy*

A small fraction of the light which interacts with tissue undergoes a scattering event that causes a change in the energy (frequency) of the incident light, a process known as inelastic scattering. Raman scattering is a specific type of inelastic scattering. In Raman scattering, the change in the energy of the incident light occurs because a portion of the energy is transferred from the photons to the molecules of the material, where it excites vibrational energy states (Stokes scattering), or, to a lesser extent, from the molecules of the material to the photons (anti-Stokes scattering)[25]. Figure 3.1 shows an energy level diagram depicting the processes by which Stokes and anti-Stokes Raman scattering occur.

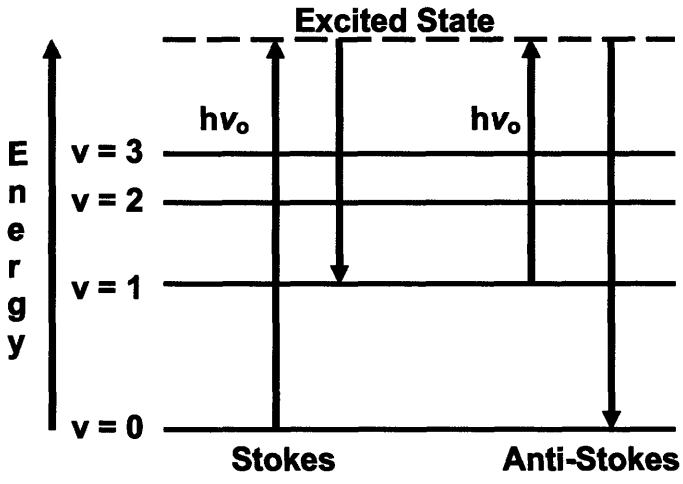

Fig 3.1 Illustration showing the processes by which Stokes Raman and anti-Stokes Raman scattering occur. In Stokes Raman scattering, the incident photon excites a molecule into a higher vibrational state excited state and only a portion of the energy is released as the molecule returns to a lower energy state. In anti-Stokes Raman scattering, a photon interacts with a molecule in a higher vibrational energy state. When the molecule transitions to a lower energy state, a portion of the absorbed energy is transferred to the photon, such that the emitted photon has more energy than the incident photon.

By analyzing a Raman spectrum, a plot of the scattered light intensity as a function of the change in energy between the incident and scattered photon in wavenumbers (Raman shift), the relative contributions of various statistical, chemical, or morphological components can be quantitatively extracted and used to develop diagnostic algorithms[26]. There are a number of Raman active biological molecules

including nucleic acids, proteins (e.g. collagen and elastin) and lipids[27]. In most tissue applications, near-infrared light is used as the excitation source in order to avoid the strong interference from tissue fluorescence at shorter wavelengths[26].

A major advantage of Raman spectroscopy is its molecular specificity. In addition, molecular vibrations are influenced by the microenvironment of functional groups, and thus this technique provides information about molecular interactions. Several considerations and challenges in the application of Raman spectroscopy to tissue diagnostics are related to the inherent weakness of the Raman signal. Ensuring an adequate signal-to-noise ratio (SNR) (which is also dependent on the tissue under study) can be difficult, and therefore powerful sources and sensitive detectors are required. The need for low or no ambient light (lights turned off) to prevent artifacts in the signal may not be possible in all clinical settings. In addition, Raman spectra require extensive processing, including the removal of the significant background fluorescence in the measured signal[16]. Despite these challenges, a number of researchers have successfully applied this technique in *in vivo* studies in tissue[28,29].

Raman Spectroscopy Studies in the Oral Cavity

Most Raman spectroscopy studies of the oral cavity have been limited to *ex vivo* frozen tissue samples, formalin-fixed samples, or animal models[30-33]. Venkatakrishna *et al.* examined 49 malignant and normal tissue samples in saline within 30 minutes of resection[34]. The tissue was excited with 785 nm light. The authors observed that the spectra from normal tissue displayed more lipid features than those from malignant samples. A total of 140 spectra were analyzed using PCA

to extract quantitative parameters and classified by applying a threshold on the Mahalanobis distance (to determine similarity and thus likelihood of membership) calculated for each sample compared to the model set of malignant spectra. Using this approach they obtained a sensitivity and specificity of 85% and 90%, respectively.

3.2.3 Optical Coherence Tomography (OCT)

Optical Coherence Tomography (OCT) is an imaging modality which measures backscattered light and provides high resolution (1-15 μm), cross-sectional images in real-time[16]. The technique, based on the Michelson interferometer, compares the backscattered light collected from the sample arm (tissue) to that from the reference arm (mirror) with a known pathlength in order to extract distance and microstructural information. Figure 3.2 shows a diagram depicting the basic components of a Michelson interferometer[35]. A light source is split into two beams, each of which is directed to one arm of the interferometer. The signal measured from the two arms recombine at the beam splitter, and the intensity of this signal is measured[16]. In OCT a low coherence light source is used therefore an interference pattern is produced only when the reflected light from both beams have traveled the same optical distance. The tomogram is generated by performing rapid axial measurements of backscattered light (by scanning the reference mirror) along successive transverse positions on the tissue[28,29,36].

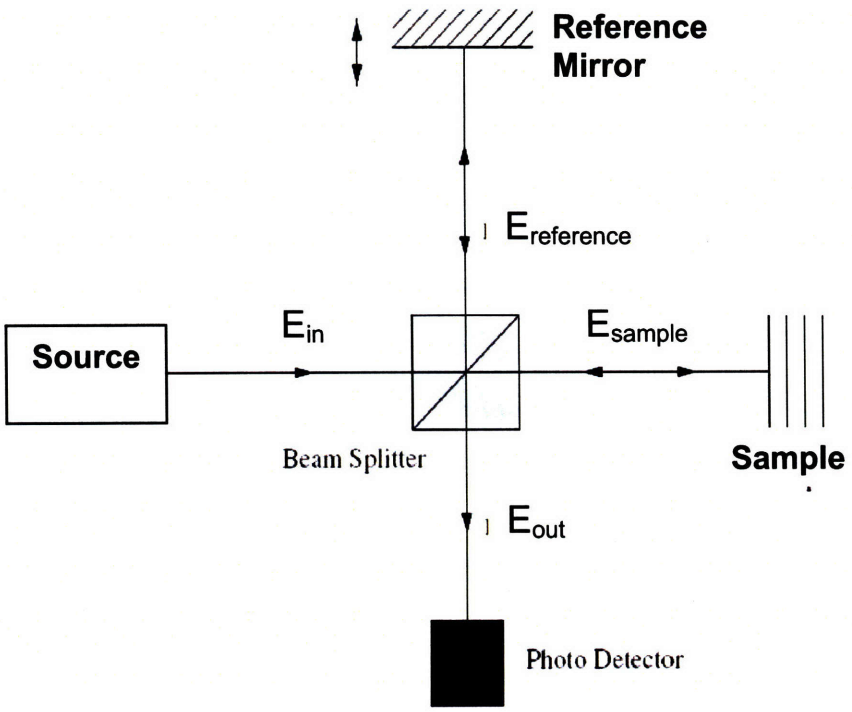

Figure 3.2 Basic components of a Michelson interferometer. The signals from the input (E_{in}), output (E_{out}), reference ($E_{reference}$) and sample (E_{sample}) arms are shown. Modified from Tomlins *et al.*

Although OCT is equivalent in principle to ultrasound cross-sectional (B-mode) imaging, it provides far greater resolution. This technique can also be applied using a fiber-based geometry, which enables tissue sites that are only accessible by an endoscope or catheter to be studied. Some limitations are that the penetration can be severely limited in the presence of significant absorption or scattering, its sensitivity to tissue movement, and the difficulty of training needed for the identification and interpretation of structures in the images.

OCT Studies in the Oral Cavity

In an early study by Feldchtein *et al.* of 5 volunteers with healthy tissue, they found marked differences between anatomic sites with non-keratinized and keratinized epithelia[37]. They noted that the presence of keratin resulted in a reduction in their ability to see deeper structures. They were able to note features such as the epithelium, lamina propria, muscle, bones, glands and vessels; however, the extent to which they were able to distinguish various features varied from site to site. In a more recent study of 41 adults, the investigators examined normal, benign, and malignant oral cavity and oropharynx tissue sites during surgical endoscopy in order to characterize normal structural components[38]. They compared the OCT images to standard images from histopathology. Using 1.31 μm light, they were able to image up to 1.6 mm in depth with a lateral and axial resolution of 15 and 10 μm, respectively. From their images, they were able to identify overlying keratin, the epithelium, the basement membrane, lamina propria, glands and vessels. Furthermore, they could observe transitions from areas of normal to cancer based on the elimination of the basement membrane and loss of normal tissue microstructures. One challenge they highlight for the application of OCT in the oral cavity is the need for the patient to remain motionless (as during a surgical procedure when patients are under general anesthesia). However, this can be overcome as the speed of data acquisition continues to increase.

3.2.4 *Fluorescence Spectroscopy*

Fluorescence arises when a molecule absorbs light with the appropriate energy to undergo an electronic transition from the ground state, dissipates a portion of the energy through non-radiative processes, and then emits light at a lower energy. Fluorophores are highly sensitive to their surroundings and thus may be used to identify changes in the tissue microenvironment relevant to cancer development perhaps not yet identifiable by gross or microscopic inspection. Factors such as the presence of quenchers (i.e. O_2,), the solvent, the pH, and nearby chromophores with which energy transfer, complex formation, or reactions can occur, greatly impact the characteristics of the emitted light[39]. A variety of fluorophores have been identified which are native to the tissue, including molecules related to metabolism and energy transport (NADH, flavin adenine dinucleotide (FAD)), structural proteins (collagen, elastin, and keratin), amino acids (tryptophan), and porphyrins (e.g. protoporphyrin IX (PPIX), Zn protoporphyrin, coproporphyrin)[40-43]. Table 3.3 lists the excitation and emission maxima or range for a variety of tissue fluorophores[17,44].

Fluorophore	Excitation Maximum [nm]	Emission Maxima [nm]
Collagen crosslinks	325	400
Elastin crosslinks	325	400
NADH	290, 340	440, 450
FAD	450	515
Tryptophan	280	350
Porphyrins	405	630, 690

Table 3.3 Excitation and emission properties of some endogenous tissue fluorophores

Fluorescence spectroscopy provides both metabolic and morphological information about the tissue. Additionally, the use of exogenous fluorophores permits enhanced

contrast and targeting of specific molecules of interest (e.g. to monitor their dynamics or localization, or treatment purposes, as in photodynamic therapy). One disadvantage is that fluorescence emission can be affected not only by the fluorophores and their local environment, but also by the processes of absorption and scattering.

Fluorescence Spectroscopy Studies in the Oral Cavity

Fluorescence spectroscopy has been extensively investigated as a tool for oral cancer detection[12,24,30,42,45-65]. Most studies collect information about the spectral profile (wavelength dependence) of the fluorescence emission. A variety of excitations between approximately 200 and 600 nm have been utilized in these studies. The quantitative spectral parameters are usually extracted using methods such as intensity ratios, area under the curve values, principal component analysis (PCA) and neural networks. Discrimination between clinically normal healthy tissue and malignant or dysplastic/malignant has been investigated in most studies and shown to be highly successful. The ability to distinguish between benign and dysplastic/malignant lesions has been less frequently investigated[66].

A few studies have performed time-resolved measurements, in which the timing of the fluorescence decay is analyzed[67]. The advantages of this technique compared to steady-state measurements, is that fluorophores whose emission spectra overlap spectrally, may be resolved temporally, and the decay is less sensitive to variables such as the instrumentation and intensity of the excitation wavelength. In one study by Chen *et al.*, clinically normal and pre-malignant lesions in the oral cavity were excited with 410 nm light and the amplitude and decay time at 633 nm

emission were extracted[67]. They could successfully distinguish members of three groups, normal, epithelial hyperplasia and verrucous hyperplasia (a type of benign white lesion), and dysplasia, respectively, into three categories using Fisher's discriminant analysis with an accuracy of 100, 75, and 93%, respectively. One of the 2 components in the exponential decay model was hypothesized to be PPIX, however the source of shorter lifetime component could not be definitively assigned based on the available literature. The authors note that the technique was time-consuming, taking approximately 30 s per measurement in order to acquire a sufficient number of photons.

3.2.5 Microscopy Techniques

Multi-photon Microscopy (MPM)

Multi-photon microscopy in a nonlinear technique in which fluorescence is generated by near-simultaneous absorption of two or more low energy photons that sum to the energy needed to excite an electronic transition[68]. The use of longer wavelength (infrared) light to excite fluorescence permits deeper light penetration, by reducing the probability of scattering and absorption. Unlike traditional 1 photon fluorescence, pulsed laser light (femtoseconds-attoseconds in between the two pulses in two-photon excitation) is used and the spot is focused to a small volume to increase the intensity and likelihood of a multi-photon event. This beam is then scanned to image the entire sample. Limiting the fluorescence to a small volume decreases photobleaching, and increases axial depth discrimination as compared to 1 photon fluorescence techniques[16]. Photodamage can occur, however, in the focal volume.

Another advantage is that the excitation spectrum is entirely resolved from the emission spectrum, such that the former can be completely filtered without substantial loss in the measured intensity. One disadvantage is that the quality of the image (contrast and resolution) can be greatly diminished if there is significant scattering of the fluorescence excitation or emission light, as occurs in tissue[16]. MPM has not been studied *in vivo* in humans for oral cancer detection, but has been applied in animal models for oral cancer[69,70].

Confocal Reflectance Microscopy (CRM)

Confocal reflectance microscopy (CRM) examines the backscattered light from tissue captured through an aperture which matches that of the delivery light, enabling a plane of focus to be selectively imaged. Some disadvantages of confocal microscopy include the small field view and limited penetration depth.

An *in vivo* confocal reflectance microscopy study by White *et al.* examined healthy mucosa from the lip and tongue with a video-rate CRM system[71]. The lateral resolution of the system was 0.5-1.0 μm and the axial resolution 3-5μm. They achieved imaging depths up to 490 and 250 μm in the lip and tongue, respectively. Imaging depth was limited by superficial scattering, especially for keratinized sites. Their system generated horizontal (*en face*) images, as compared to the normal transverse view of histopathology. Only anterior portions of the oral cavity could be assessed because of the need for a bulky tissue stabilizer to minimize motion artifacts during their measurements, which took several seconds per site. Their images permitted greater cellular detail to be appreciated than that of a corresponding

histological section taken from the same site. Interestingly, they calculated several quantitative parameters related to the tissue architecture (i.e. epithelial thickness, cellular and nuclear diameters of epithelial cells) and found statistically significant differences when these values were compared to the same metrics calculated from histological sections, which may be due to artifacts during processing of tissue specimens.

3.3 Spectroscopic Models

In the present study, we focus on the detection of oral cancer using two spectroscopic modalities: Diffuse Reflectance Spectroscopy (DRS) and Intrinsic Fluorescence Spectroscopy (IFS). Our laboratory has developed physical models which we apply to the measured spectral data in order to extract quantitative parameters. In this section we provide a detailed description of these models. Light Scattering Spectroscopy (LSS) is a third modality, which was also developed by our laboratory and previously used in the analysis of clinical data. In the course of the present work, new findings related to this technique were discovered and we discuss why this technique is no longer applied to the clinical data.

3.3.1 Diffuse Reflectance Spectroscopy

Most models describing the propagation of scattered light in tissue are based on the steady-state diffusion approximation to the radiative transfer equation[72]. The

radiative transfer equation describes the transport of energy through a medium containing particles and has been applied to a variety of scattering phenomena[72,73]. The radiative transport equation treats light as a collection of photons and neglects the field properties of light. While the radiative transport equation can only be solved numerically (there is no closed form solution), the diffusion equation can be solved analytically. The diffusion approximation assumes that light transport is dominated by scattering, as is often the case for tissue, such that photons are scattered uniformly in all directions and engage in what can be thought of as a random walk through the medium.

Diffuse Reflectance Models

In the diffusion-based description of reflectance there are two key parameters for describing the interaction of light with scatterers and absorbers: the reduced scattering coefficient (μ_s') and absorption coefficient (μ_a), respectively. These parameters are wavelength dependent and the inverse of these values represents the average distance (step-size) traveled by a photon before encountering a scatterer/absorber. The reduced scattering coefficient is calculated as follows:

$$\mu_s'(\lambda) = \mu_s \cdot (1-g), \quad [\text{mm}^{-1}] \tag{3.1}$$

where μ_s is the scattering coefficient, and g, is the anisotropy parameter. Because light is not isotropically scattered in tissue (an equal probability of scattering in all directions), but instead is strongly forward-directed, the anisotropy parameter, g, applies a correction enabling the scattering to be treated in a manner equivalent to the isotropic case, which is one of the requirements for the diffusion approximation. The

anisotropy parameter is wavelength-dependent and represents the average cosine of the scattering angle. An equation for spatially-resolved diffuse reflectance in the radial direction, $R_d(r)$, has been derived by Farrell *et al.* based on this approximation[74]:

$$R_d(r) = \frac{a'}{4\pi}\left[\frac{1}{\mu_t'}\cdot\left(\mu_{eff}+\frac{1}{r_1}\right)\cdot\frac{e^{-\mu_{eff}r_1}}{r_1^2}+\left(\frac{1}{\mu_t'}+2z_b\right)\cdot\left(\mu_{eff}+\frac{1}{r_2}\right)\cdot\frac{e^{-\mu_{eff}r_2}}{r_2^2}\right], \quad (3.2)$$

where,

$$r_1 = [z_0^2 + r^2]^2,$$
$$r_2 = [(z_0 + 2z_b)^2 + r^2]^{1/2},$$
$$a' = \frac{\mu_s'}{\left[\mu_s' + \mu_a\right]},$$
$$\mu_{eff} = \frac{1}{3\mu_a\left[\mu_s' + \mu_a\right]},$$
$$\mu_t' = \mu_a + \mu_s',$$
$$z_b = 2AD,$$
$$A = \frac{1+y}{1-y},$$
$$y = -1.44n_{rel}^{-2} + 0.710n_{rel}^{-1} + 0.0636n_{rel} + 0.668.$$

The quantities r_1 and r_2 are related to the boundary conditions, which specify that the fluence at the interface must be zero. The term z_o is equal to $1/\mu_t'$. The quantity a' is the transport albedo and μ_{eff} denotes the effective attenuation coefficient. To account for the mismatch in refractive index at the tissue-surface interface, the parameter A describes the internal reflection at the interface and is calculated using y as described. The term n_{rel} is the relative refractive index. Similar to Fick's 1st Law of Diffusion, the diffuse photons are characterized by a diffusion coefficient, D, where

$$D = \frac{1}{3\left[\mu_s' + \mu_a\right]} \quad [\text{mm}]. \tag{3.3}$$

In the case where light is delivered to and collected from the tissue via an optical fiber probe, R_d must be integrated over the area of the delivery and collection fibers, respectively. The model of Farrell *et al.* assumes a semi-infinite medium and describes the diffuse reflectance collected at the tissue surface for a single scatterer source.

The values for the scattering and absorption coefficients can be extracted by modeling the reflectance data. With an increasing number of scattering/absorption events within a medium, each parameter increases in value. The diffusion approximation is valid when $\mu_s' >> \mu_a$ and breaks down in highly absorbing samples or very dilute samples, in which photons do not sample the medium to any great extent, and thus are not sufficiently randomized. This model also becomes inaccurate at small source-detector separations, where the collected signal represents light which has traveled only very short distances within the tissue and has not experienced sufficient scattering events to become diffusive.

Our laboratory has developed a simple analytical DRS model, based on the model of Farrell *et al.*, which is specific for our probe configuration[13]. The diffuse reflectance is once again characterized by the reduced scattering coefficient and absorption coefficient. The tissue is considered to be a homogenous, semi-infinite medium in which the re-emitted signal corresponds to the diffusely scattered light which escapes after losses in intensity due to absorption. The equation developed by

Zonios *et al.* for the wavelength-dependent diffuse reflectance collected from the probe, R_{meas}, is expressed as follows:

$$R_{meas}(\lambda) = \frac{\mu_s'}{\mu_s' + \mu_a}\left\{e^{-\mu z_0} + e^{-(1+\frac{4}{3}A)\cdot\mu z_0} - z_0\frac{e^{-\mu r_1'}}{r_1'} - (1+\frac{4}{3}A)\cdot z_0\frac{e^{-\mu r_2'}}{r_2'}\right\} \quad (3.4)$$

where,

$z_0 = 1/(\mu_s' + \mu_a)$
$r_1' = [(z_0^2 + r_c'^2]^{1/2}$, and
$r_2' = [z_0^2(1+(4/3)\cdot A)^2 + r_c'^2]^{1/2}$.

The properties of the probe are taken into account, as in the full model described by Farrell *et al.*, but a few additional simplifications are included. First, the delivery radius of the probe, r_d, is neglected (treated as a point source). Second, the collection radius of the probe, r_c, is modeled by a single term, r_c', which corresponds to an effective radius of collection. This latter term is an empirical value derived by calibrating the probe with a phantom, a simple tissue model with known scattering and absorption properties. Based on a calibration of our probe using a solution of beads with known reflectance properties, we use a value of 0.36 mm for r_c' in our analysis (the reduced scattering coefficient and absorption coefficient are expressed in units of mm^{-1}). These assumptions reduce the double integral in the original form of the equation to a single integral with a simple analytical solution. We use a value of 1.4 for n_{rel}, so that A = 3.2[13].

Modeling the Reduced Scattering Coefficient

In our study of oral tissue, the reduced scattering coefficient was constrained to follow a double power law, such that

$$\mu_s'(\lambda) = A \cdot \left(\frac{\lambda}{\lambda_o}\right)^{-B} + C \cdot \left(\frac{\lambda}{\lambda_o}\right)^{-4} . [\text{mm}^{-1}] \quad (3.5)$$

Here, the wavelength, λ, is expressed in units of microns, and $\lambda_0 = 1$ μm. This power law description (with and without the second term) has been widely used to model the wavelength-dependence of the scattering coefficient parameter in studies of both cells and tissue[75-78]. From modeling this parameter we extract three scattering parameters that describe the tissue, A, B, and C[76,77].

The A parameter is a scaling parameter related to the overall magnitude of scattering. In a study of skin, the collagen fibers of the dermis were shown to be the major source of light scattering- contributing mostly to scattering from larger particles (Mie scattering), and to a lesser extent, scattering by small particles (Rayleigh scattering)[79]. Mourant *et al.* showed that when modeling scattering from suspended cells with a Gaussian distribution, the exponent of the power law was sensitive to the standard deviation (width of the size distribution) and increased until reaching a value of 4 (the Raleigh limit) as the size of the scatterers decreased. They also found that the exponent is particularly sensitive to the smaller particles in the distribution[75]. The C term represents the magnitude of scattering by Rayleigh particles.

Modeling the Absorption Coefficient

The absorption coefficient was modeled as the sum of the contributions from two absorbers, hemoglobin and β-carotene:

$$\mu_a(\lambda) = \mu_a^{Hb}(\lambda) + \mu_a^{\beta C}(\lambda). \text{ [mm}^{-1}\text{]} \tag{3.6}$$

Plots of the extinction spectra for Hb, HbO_2, and β-carotene are shown in Figure 3.3.

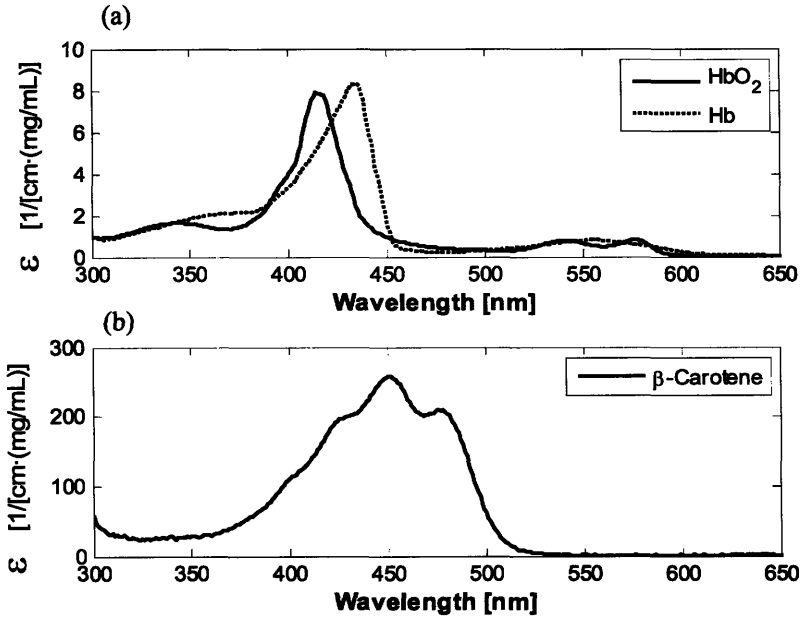

Figure 3.3 A plot of the extinction spectrum, ε, for three absorbers found in oral tissue: (a) Hb (dotted black line) and HbO_2 (solid black line) (b) β-Carotene.

Hemoglobin absorption was modeled using a correction for vessel packaging previously described in the literature[80-82]. In our analysis, we adopt the formulation of the model developed by Svaasand *et al.*, and also presented by van Veen *et al*[80,82]. Vessel packaging accounts for alterations in the shape and intensity of the hemoglobin absorption peaks as a result of the absorber being confined to discrete finite-sized blood vessels, rather than homogeneously distributed throughout the tissue. This model assumes that there is a single-sized vessel distributed uniformly

throughout the tissue. We further impose a lower limit on the fitted vessel radius of 2.5 μm, since the minimum diameter of a capillary is on the order of 5-7 μm[83].

Using this model, the absorption coefficient of hemoglobin is represented by the product of the absorption coefficient of hemoglobin in whole blood, μ_a^{blood}, a wavelength-dependent correction factor which alters the shape of the absorption spectrum, K, and the volume fraction of blood sampled, ν. The absorption coefficient for hemoglobin was described as follows:

$$\mu_a^{Hb}(\lambda) = K(\lambda) \cdot \nu \cdot \mu_a^{blood}(\lambda) \quad [\text{mm}^{-1}]. \tag{3.7}$$

The correction factor K is calculated by the following equation,

$$K(\lambda) = \left\{ \frac{1 - \exp[-2 \cdot \mu_a^{blood}(\lambda) \cdot R\ ]}{2 \cdot R \cdot \mu_a^{blood}(\lambda)} \right\}, \tag{3.8}$$

where R is the effective vessel radius in units of millimeters and the absorption coefficient of blood, $\mu_a^{blood}(\lambda)$ is in units of mm^{-1}. The volume fraction, ν, is described by the following ratio:

$$\nu = \text{cHb}/150, \tag{3.9}$$

where cHb is the extracted concentration of hemoglobin in units of mg/mL. The concentration of hemoglobin in whole blood was assigned a value of 150 mg/mL. The equation for μ_a^{blood} is given by

$$\mu_a^{blood}(\lambda) = 0.1 \cdot \ln(10) \cdot 150[\alpha \cdot \varepsilon_{HbO_2}(\lambda) + (1-\alpha) \cdot \varepsilon_{Hb}(\lambda)] \ , [\text{mm}^{-1}] \qquad (3.10)$$

where α is the oxygen saturation, ε is the extinction coefficient in units of 1/[cm·(mg/mL)], HbO_2 denotes oxyhemoglobin, and Hb denotes deoxyhemoglobin[19]. The dramatic effect of the correction factor, *K*, on the lineshape and intensity of the absorption coefficient of hemoglobin is shown in Figure 3.4.

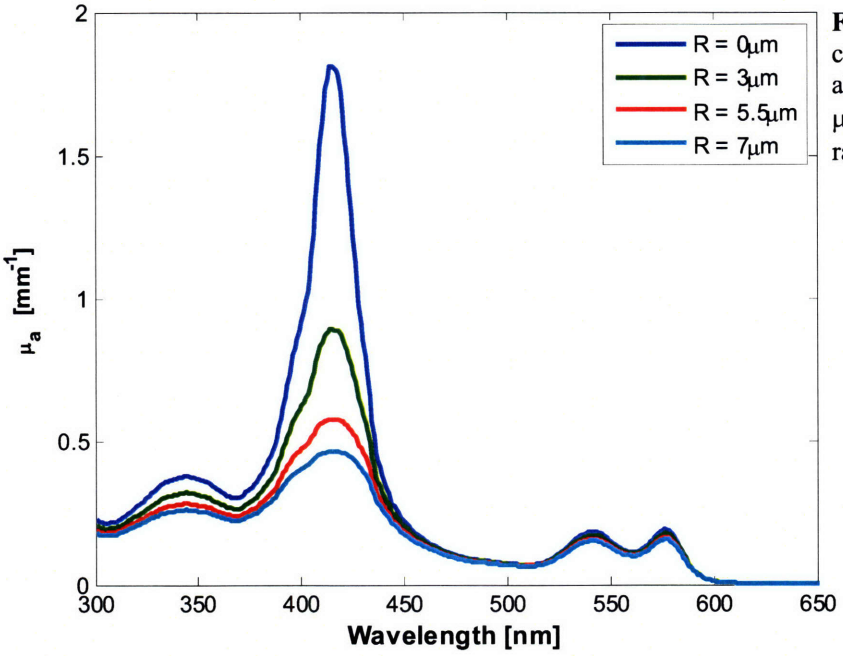

Figure 3.4 A plot of the absorption coefficient of hemoglobin in the absence of vessel packaging R= 0 μm and for three different vessel radii: 3, 5.5, and 7 μm, respectively.

Figure 3.5 shows synthetic diffuse reflectance curves generated with the model of Zonios *et al.* Parameters consistent with those obtained from clinical data collected from the oral cavity were used to calculate μ_s' and μ_a. Figure 3.5(a) shows the effect of varying cHb while the vessel radius, R, is held constant. The plot shows that the depth of the 420 nm and 540/580 nm hemoglobin peaks increases with increasing

hemoglobin concentration. Figure 3.5(b) shows the effect of varying the vessel radius, R, while cHb is held constant. In this case, the intensity of the 420 nm absorption peak decreases relative to that of the 540/580 nm peaks, while the latter peaks remain essentially unchanged.

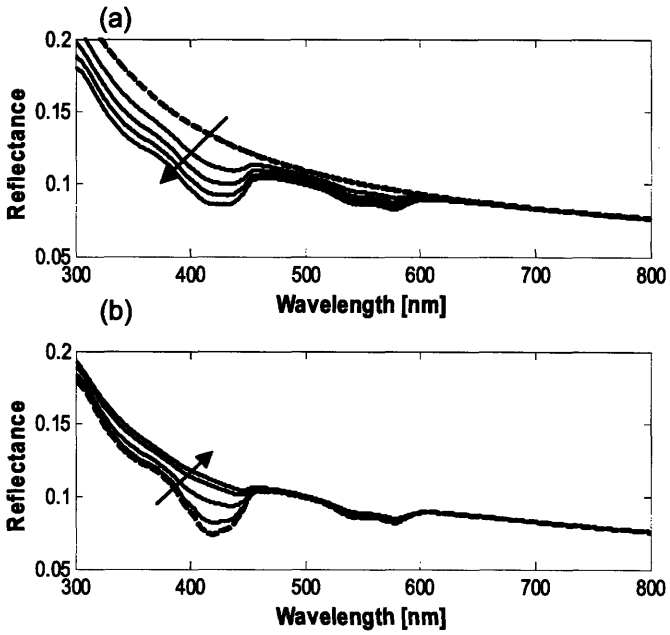

Figure 3.5 Synthetic diffuse reflectance data generated using the model of Zonios *et al.* based on scattering and absorption parameters consistent with those obtained from clinical data. In each plot, the dashed curve represents the curve for the minimum value of the varied parameter and the arrow represents the direction of increasing value for the varied parameter. (A). Reflectance curves generated for hemoglobin concentrations varying between 0 and 2 mg/mL, while the vessel radius was held constant. (B) Reflectance curves generated for vessel radii between 2.5 and 20 μm, while the hemoglobin concentration was held constant.

The absorption coefficient for β-carotene was modeled as follows:

$$\mu_a^{\beta C}(\lambda) = 0.1 \cdot \ln(10) \cdot c\beta C \cdot \varepsilon_{\beta C}(\lambda), [\text{mm}^{-1}] \tag{3.11}$$

where cβC denotes the concentration of β-carotene in units of mg/mL, and $\varepsilon_{\beta C}$ is the extinction coefficient β-carotene in units of 1/[cm·(mg/mL)][18]. The complete description of the absorption coefficient was as follows:

$$\mu_a(\lambda) = K(\lambda) \cdot \nu \cdot \mu_a^{blood}(\lambda) + 0.1 \cdot \ln(10) \cdot c\beta C \cdot \varepsilon_{\beta Car}(\lambda) . [\text{mm}^{-1}] \tag{3.12}$$

3.3.2 *Intrinsic Fluorescence Spectroscopy*

Intrinsic fluorescence spectroscopy (IFS) is a technique that was developed by our laboratory to extract the fluorescence emitted by molecules native to the tissue (autofluorescence) from the measured fluorescence, without alterations in intensity and lineshape due to scattering and absorption[84,85]. Intrinsic fluorescence, unlike the measured fluorescence, represents a linear combination of the individual fluorophores present within the tissue.

Fluorescence photons undergo scattering and absorption between the path of origin and detection, which can cause significant distortions in the detected lineshape and intensity. These distortions prevent the quantitative extraction of the specific biochemical tissue components. Because fluorescence and reflectance photons undergo approximately the same scattering and absorption events, the alterations in the measured fluorescence can be accounted for and corrected, using a reflectance spectrum that is collected from the same location. Extracting the intrinsic fluorescence is particularly important for ensuring that changes in absorption (in particular due to hemoglobin) and scattering do not affect the fluorescence parameters and lead to misinterpretation of the data. The intrinsic fluorescence, f_{xm}, can be calculated as follows:

$$f_{xm} = \frac{F_{xm}}{\frac{1}{\mu_{sx} l}\left(\frac{R_{ox}R_{om}}{\varepsilon_x \varepsilon_m}\right)^{1/2}\left(\frac{R_x}{R_{ox}}\right)\left(\frac{R_m}{R_{om}} + \varepsilon_m\right)} \qquad (3.13)$$

$$\varepsilon = e^{\beta} - 1$$

$$\beta = S(1-g)$$

The subscripts x and m refer to the dependence of the quantities on the excitation and emission wavelengths, respectively. F_{xm} is the measured fluorescence and R_o is the measured reflectance in the absence of absorption. The constant *l*, has units of length, and is related to the effective depth from which the fluorescence is collected by the probe and is determined by the medium, as well as the delivery-collection geometry of the probe. The parameter *S* represents a probe-specific constant, and g is the anisotropy parameter. In our analysis, we use values obtained from studies of skin as a reasonable approximation for *g*.

This method was developed based on the photon migration model, which describes photons traveling through a turbid medium along paths with discrete nodes, at which a scattering or absorption event occurs[85,86]. At each node, a portion of the light is lost due to absorption and a fraction is scattered and continues to travel through the medium. In this picture, the reflectance is related to the product of the fraction of unabsorbed photons at each interaction event (albedo), and the probability that a photon entering the medium and undergoing *y* scattering events will escape and be collected. An analytical expression for the escape probability, ρ_n, was determined based on Monte Carlo simulations by Wu *et al.* and modified by Zhang *et al.*, such that the reflectance, *R*, in the absence of absorption can de described as follows:

$$R(\lambda) = \sum_{y=1}^{\infty} a^n \rho_n = \frac{\kappa \cdot a \cdot \exp(-\beta)}{1 - a \cdot \exp(-\beta)}, \tag{3.14}$$

where

$$\rho_n = a \cdot \exp(-\beta n)$$

$$\beta = S(1-g),$$
$$\kappa = \varepsilon \cdot R(\lambda),$$
$$\varepsilon = \exp(\beta) - 1,$$
$$a = \mu_s/(\mu_s + \mu_a)$$

The parameter S is a constant related to the probe, g is the anisotropy parameter, and a is the albedo.

To extend the picture to incorporate fluorescence, the fluorescence is considered to occur at the $(i+1)^{th}$ node, after an initial i scattering events at the excitation wavelength. After being absorbed by the fluorophore, the fluorescence photon undergoes y-i-1 scattering events at the emission wavelength, where y is the total number of scattering events. The measured fluorescence can then be defined in a similar manner as the reflectance as follows:

$$F_{xm} \equiv \sum_{y=1}^{\infty} \sum_{i=0}^{y-1} a_x^{\,i+1} \left(\frac{f_{xm}}{\mu_{sx} \cdot l} \right) a_m^{\,y-i-1} \rho_{ni}, \tag{3.15}$$

where

$$f_{xm} = \left(\frac{I_x}{h\nu_x} \right) \mu_{fx} \cdot l \cdot \phi_{xm} h\nu_m, \tag{3.16}$$

$$\rho_{ni} = (a_x a_m)^{1/2} \cdot \exp[-\beta_x (i+1) \exp[-\beta_m (y-i-1)],$$

The intensity is denoted by I and the absorption coefficient of the fluorophore as μ_f. The energy per the photon is calculated as the product of Planck's constant, h, and the frequency of the light, ν. The quantum yield of the fluorophore is denoted by ϕ, and l,

as before is a probe-specific constant. The quantity f_{xm} is defined as the intrinsic fluorescence. By inserting the terms ρ_n and ρ_{ni} into the equations for $R(\lambda)$ and F_{xm}, we obtain the formula for the intrinsic fluorescence derived above (Eq. 3.13). This model assumes a homogeneous medium and does not account for subsequent secondary absorption of fluorescence that occurs within the medium, although these are not expected to be significant effects.

In order to apply the IFS model, we need to compute the reflectance in the absence of absorption at the excitation and emission wavelengths, R_{ox} and R_{om}, respectively. To do this we first used the scattering parameters extracted from the modeled fit to the reflectance spectrum to calculate the wavelength-dependence of μ_s' (Eq. 3.5). R_{ox} and R_{om} are then calculated by setting $\mu_a(\lambda)$ to zero, inputting $\mu_s'(\lambda)$ into the model of Zonios *et al.* (Eq. 3.4), and evaluating the equation either at each of the ten excitations wavelengths (R_{ox}), or for the entire range of emission wavelengths (R_{om}). The R_x term in the IFS model is calculated by evaluating the equation of Zonios *et al.*, using the hemoglobin concentration extracted from the fit to the diffuse reflectance data to calculate the input $\mu_a(\lambda)$. R_m was obtained directly from the calibrated reflectance. Based on experiments with tissue phantoms, the constant S was set to 1 and the constant l was set to 220 μm.

3.3.3 Light Scattering Spectroscopy

Light scattering spectroscopy (LSS) studies the spectral variation of singly-scattered elastic light from the most superficial layers of tissue. This signal is attributed to scattering from epithelial nuclei. The reflectance signal collected from

tissue represents both singly and diffusely scattered light. Our laboratory developed a novel method for separating the much weaker singly-scattered light signal from the large diffusive background using a model-based approach[87,88]. In this approach, the light scattering signal is obtained by subtracting the modeled diffuse scattering component (DRS fit, Section 3.3.1) from the measured reflectance. By modeling the amplitude and frequency of this residual signal, which corresponds to the light scattering signal, the nuclear density and size distribution were then extracted.

After an in depth investigation of this model, a number of findings were discovered that provided strong evidence that this technique is not reliable. While the initial studies and development of this technique were based on work with cells, it was found that in tissue, imperfect modeling of hemoglobin absorption using the DRS model resulted in hemoglobin absorption features in the residual signal. This was particularly the case when vessel packaging was not taken into account in the DRS model. Figure 3.6 shows an example of a reflectance spectrum collected from the oral cavity, and the modeled fit with the DRS model either with or without taking vessel packaging into account [Fig. 3.6(a)]. The LSS residual signals extracted using the data with each of the respective fits in Fig. 3.6(a) are shown in Fig. 3.6(b). The hemoglobin features are clearly visible in the LSS residual obtained from the fit in which vessel packaging was not taken into account. The imperfect fit alters both the amplitude and frequency of the oscillations.

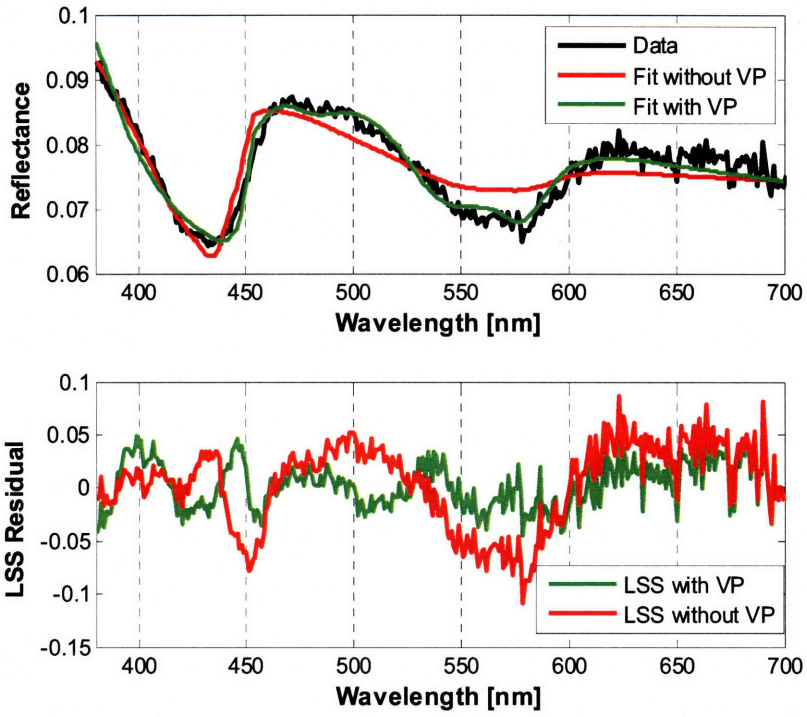

Figure 3.6 (a) A plot of a reflectance spectrum collected from the lateral surface of the tongue (black line), and the modeled fit without vessel packaging (red line), and with vessel packaging (green line). (b) The LSS residual signal obtained from the fit without vessel packaging (red line), and with vessel packaging (green line). VP: vessel packaging

The amplitude of the LSS signal was found to be on the order of 5%, as shown in Figure 3.6(b). In an analysis performed on esophageal tissue, hemoglobin features were visible and oscillatory amplitudes as great as 20% were observed[10,89]. In contrast to these results, we recently performed a numerical simulation to predict the expected amplitude of the LSS signal in tissue and it was found to be <2%[89]. Given that the typical noise in the clinical measurements is on the order of 5%, we expect this to also contribute significantly to the residual obtained after subtracting the fit to the data. Therefore, reliably extracting LSS parameters from this signal may be more difficult.

As another test, we generated numerical (synthetic) reflectance data using parameters consistent with those obtained from clinical data as inputs into the Zonios reflectance model. The scattering coefficient was calculated using Eq. 3.5 and the

absorption for hemoglobin was modeled either with vessel packaging (Eq. 3.7) or without vessel packaging (equivalent to Eq. 3.10 without the factor of 150). Random Gaussian noise was added to all the spectra in the amounts typically found in clinical data measured from the oral cavity. In the first series of reflectance spectra, vessel packaging effects were included in the model and each parameter (*A*, *B*, *C*, cHb, α, the effective vessel radius, the wavelength-dependence (shape) of the noise, and magnitude of the noise) was systematically varied, while the others were held constant. The resulting numerical reflectance data were fit over the range of 380-700 nm using the Zonios model, but without including vessel packaging in modeling the absorption. LSS parameters were extracted from the residual spectra using the published method[87]. All reflectance and residual data were fit over the range of 380-700 nm. Since the numerically generated data was fit without vessel packaging, hemoglobin absorption features were clearly visible in the residual spectra. If we apply the model of Perelman *et al.*, we extract values for the mean diameter (~11-14 μm) and standard deviations of the nuclei (~0-5 μm) that are consistent with those obtained from pathology even in the absence of nuclear scattering. Furthermore, the extracted LSS parameters were significantly affected by the wavelength range used in the LSS analysis, the concentration of hemoglobin present, and the magnitude and wavelength-dependence (shape) of the noise. For example, when the analysis range was changed from 350-700 nm to 390-700 nm, the values for the extracted mean diameter and the percentage of enlarged nuclei shifted by 3 μm and 40%, respectively.

Previous studies in our laboratory combined all three modalities, DRS, LSS, and IFS, a technique known as Tri-modal spectroscopy (TMS)[11,12]. However, from the compelling evidence mentioned above, it was decided that the LSS technique was not a robust method and the signal generated using the model-based approach was unlikely to be due to epithelial nuclei. Therefore only DRS and IFS were used in the present work.

3.4 References

1. M. M. Mueller and N. E. Fusenig, "Tumor-stroma interactions directing phenotype and progression of epithelial skin tumor cells," *Differentiation*, **70**(9-10), 486-497 (2002).
2. R. C. K. Jordan, M. Macabeo-Ong, C. H. Shiboski, N. Dekker, D. G. Ginzinger, D. T. W. Wong and B. L. Schmidt, "Overexpression of matrix metalloproteinase-1 and-9 mRNA is associated with progression of oral dysplasia to cancer," *Clinical Cancer Research*, **10**(19), 6460-6465 (2004).
3. K. I. Tosios, N. Kapranos and S. I. Papanicolaou, "Loss of basement membrane components laminin and type IV collagen parallels the progression of oral epithelial neoplasia," *Histopathology*, **33**(3), 261-268 (1998).
4. S. Warnakulasuriya, "Histological grading of oral epithelial dysplasia: revisited," *Journal of Pathology*, **194**(294-297 (2001).
5. I. vanderWaal, K. P. Schepman, E. H. vanderMeij and L. E. Smeele, "Oral leukoplakia: a clinicopathological review," *Oral Oncology*, **33**(5), 291-301 (1997).
6. R. S. Cotran, V. Kumar, T. Collins and S. L. Robbins, *Robbins pathologic basis of disease*, Saunders, Philadelphia, 1999
7. M. Macluskey, L. M. Chandrachud, S. Pazouki, M. Green, D. M. Chisholm, G. R. Ogden, S. L. Schor and A. M. Schor, "Apoptosis, proliferation, and angiogenesis in oral tissues. Possible relevance to tumour progression," *Journal of Pathology*, **191**(4), 368-375 (2000).
8. D. Arifler, I. Pavlova, A. Gillenwater and R. Richards-Kortum, "Light scattering from collagen fiber networks: Micro-optical properties of normal and neoplastic stroma," *Biophysical Journal*, **92**(9), 3260-3274 (2007).
9. A. Amelink, O. P. Kaspers, H. Sterenborg, J. E. van der Wal, J. L. N. Roodenburg and M. J. H. Witjes, "Non-invasive measurement of the morphology and physiology of oral mucosa by use of optical spectroscopy," *Oral Oncology*, **44**(1), 65-71 (2008).
10. I. Georgakoudi, B. C. Jacobson, J. Van Dam, V. Backman, M. B. Wallace, M. G. Muller, Q. Zhang, K. Badizadegan, D. Sun, G. A. Thomas, L. T. Perelman and M. S. Feld, "Fluorescence, reflectance, and light-scattering spectroscopy for evaluating dysplasia in patients with Barrett's esophagus," *Gastroenterology*, **120**(7), 1620-1629 (2001).
11. I. Georgakoudi, E. E. Sheets, M. G. Muller, V. Backman, C. P. Crum, K. Badizadegan, R. R. Dasari and M. S. Feld, "Trimodal spectroscopy for the detection and characterization of cervical precancers in vivo," *American Journal of Obstetrics and Gynecology*, **186**(3), 374-382 (2002).
12. M. G. Muller, T. A. Valdez, I. Georgakoudi, V. Backman, C. Fuentes, S. Kabani, N. Laver, Z. M. Wang, C. W. Boone, R. R. Dasari, S. M. Shapshay and M. S. Feld, "Spectroscopic detection and evaluation of morphologic and biochemical changes in early human oral carcinoma," *Cancer*, **97**(7), 1681-1692 (2003).
13. G. Zonios, L. T. Perelman, V. M. Backman, R. Manoharan, M. Fitzmaurice, J. Van Dam and M. S. Feld, "Diffuse reflectance spectroscopy of human

adenomatous colon polyps in vivo," *Applied Optics*, **38**(31), 6628-6637 (1999).

14. M. C. Skala, K. M. Riching, A. Gendron-Fitzpatrick, J. Eickhoff, K. W. Eliceiri, J. G. White and N. Ramanujam, "In vivo multiphoton microscopy of NADH and FAD redox states, fluorescence lifetimes, and cellular morphology in precancerous epithelia," *Proceedings of the National Academy of Sciences of the United States of America*, **104**(49), 19494-19499 (2007).
15. I. Georgakoudi, B. C. Jacobson, M. G. Muller, E. E. Sheets, K. Badizadegan, D. L. Carr-Locke, C. P. Crum, C. W. Boone, R. R. Dasari, J. Van Dam and M. S. Feld, "NAD(P)H and collagen as in vivo quantitative fluorescent biomarkers of epithelial precancerous changes," *Cancer Research*, **62**(3), 682-687 (2002).
16. T. Vo-Dinh, *Biomedical photonics handbook*, CRC Press, Boca Raton, Fla., 2003
17. R. Richards-Kortum and E. SevickMuraca, "Quantitative optical spectroscopy for tissue diagnosis," *Annual Review of Physical Chemistry*, **47**(555-606 (1996).
18. H. Du, R. C. A. Fuh, J. Z. Li, L. A. Corkan and J. S. Lindsey, "PhotochemCAD: A computer-aided design and research tool in photochemistry," *Photochemistry and Photobiology*, **68**(2), 141-142 (1998).
19. S. A. Prahl, "Tabulated Molar Extinction Coefficient for Hemoglobin in Water," Oregon Medical Laser Center http://omlc.ogi.edu/spectra/hemoglobin/summary.html, Accessed: March 8, 2008
20. V. V. Tuchin, *Tissue optics : light scattering methods and instruments for medical diagnosis, Tutorial texts in optical engineering ; TT 38*, SPIE Optical Engineering Press, Bellingham, Wash., 2000
21. A. Sharwani, W. Jerjes, V. Salih, B. Swinson, I. J. Bigio, M. El-Maaytah and C. Hopper, "Assessment of oral premalignancy using elastic scattering spectroscopy," **42**(4), 343-349 (2006).
22. D. Hattery, B. Hattery, V. Chernomordik, P. Smith, M. Loew, J. Mulshine and A. Gandjbakhche, "Differential oblique angle spectroscopy of the oral epithelium," **9**(5), 951-960 (2004).
23. N. Subhash, J. R. Mallia, S. S. Thomas, A. Mathews and P. Sebastian, "Oral cancer detection using diffuse reflectance spectral ratio R540/R575 of oxygenated hemoglobin bands," *Journal of Biomedical Optics*, **11**(1), (2006).
24. D. C. G. de Veld, M. Skurichina, M. J. H. Wities, R. P. W. Duin, H. J. C. M. Sterenborg and J. L. N. Roodenburg, "Autofluorescence and diffuse reflectance spectroscopy for oral oncology," *Lasers in Surgery and Medicine*, **36**(5), 356-364 (2005).
25. V. V. Tuchin, *Handbook of optical biomedical diagnostics*, SPIE Press, Bellingham, Wash., USA, 2002
26. E. B. Hanlon, R. Manoharan, T. W. Koo, K. E. Shafer, J. T. Motz, M. Fitzmaurice, J. R. Kramer, I. Itzkan, R. R. Dasari and M. S. Feld, "Prospects for in vivo Raman spectroscopy," *Physics in Medicine and Biology*, **45**(2), R1-R59 (2000).

27. N. K. Chaudhury, S. Chandra and T. L. Mathew, "Oncologic applications of biophotonics - Prospects and problems," *Applied Biochemistry and Biotechnology*, **96**(1-3), 183-204 (2001).
28. J. T. Motz, M. Fitzmaurice, A. Miller, S. J. Gandhi, A. S. Haka, L. H. Galindo, R. R. Dasari, J. R. Kramer and M. S. Feld, "In vivo Raman spectral pathology of human atherosclerosis and vulnerable plaque," *Journal of Biomedical Optics*, **11**(2), - (2006).
29. A. Molckovsky, L. M. W. K. Song, M. G. Shim, N. E. Marcon and B. C. Wilson, "Diagnostic potential of near-infrared Raman spectroscopy in the colon: differentiating adenomatous from hyperplastic polyps," *Gastrointestinal Endoscopy*, **57**(3), 396-402 (2003).
30. D. C. G. de Veld, T. C. B. Schut, M. Skurichina, M. J. H. Witjes, J. E. Van der Wal, J. L. N. Roodenburg and H. J. C. M. Sterenborg, "Autofluorescence and Raman microspectroscopy of tissue sections of oral lesions," *Lasers in Medical Science*, **19**(4), 203-209 (2005).
31. R. Malini, K. Venkatakrishna, J. Kurien, K. M. Pai, L. Rao, V. B. Kartha and C. M. Krishna, "Discrimination of normal, inflammatory, premalignant, and malignant oral tissue: A Raman spectroscopy study," *Biopolymers*, **81**(3), 179-193 (2006).
32. A. P. Oliveira, R. A. Bitar, L. Silveira, R. A. Zangaro and A. A. Martin, "Near-infrared Raman spectroscopy for oral carcinoma diagnosis," *Photomedicine and Laser Surgery*, **24**(3), 348-353 (2006).
33. C. M. Krishna, G. D. Sockalingum, J. Kurien, L. Rao, L. Venteo, M. Pluot, M. Manfait and V. B. Kartha, "Micro-Raman spectroscopy for optical pathology of oral squamous cell carcinoma," *Applied Spectroscopy*, **58**(9), 1128-1135 (2004).
34. K. Venkatakrishna, J. Kurien, K. M. Pai, M. Valiathan, N. N. Kumar, C. M. Krishna, G. Ullas and V. B. Kartha, "Optical pathology of oral tissue: A Raman spectroscopy diagnostic method," *Current Science*, **80**(5), 665-669 (2001).
35. P. H. Tomlins and R. K. Wang, "Theory, developments and applications of optical coherence tomography," *Journal of Physics D-Applied Physics*, **38**(15), 2519-2535 (2005).
36. A. S. Haka, Z. Volynskaya, J. A. Gardecki, J. Nazemi, J. Lyons, D. Hicks, M. Fitzmaurice, R. R. Dasari, J. P. Crowe and M. S. Feld, "In vivo margin assessment during partial mastectomy breast surgery using Raman spectroscopy," *Cancer Research*, **66**(6), 3317-3322 (2006).
37. F. I. Feldchtein, G. V. Gelikonov, V. M. Gelikonov, R. R. Iksanov, R. V. Kuranov, A. M. Sergeev, N. D. Gladkova, M. N. Ourutina, J. A. Warren and D. H. Reitze, "In vivo OCT imaging of hard and soft tissue of the oral cavity," *Optics Express*, **3**(6), 239-250 (1998).
38. J. M. Ridgway, W. B. Armstrong, S. Guo, U. Mahmood, J. P. Su, R. P. Jackson, T. Shibuya, R. L. Crumley, M. Gu, Z. P. Chen and B. J. F. Wong, "In vivo optical coherence tomography of the human oral cavity and oropharynx," *Archives of Otolaryngology-Head & Neck Surgery*, **132**(10), 1074-1081 (2006).

39. J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*, Kluwer Academics/Plenum Publishers, New York, 1999
40. S. K. Chang, N. Marin, M. Follen and R. Richards-Kortum, "Model-based analysis of clinical fluorescence spectroscopy for in vivo detection of cervical intraepithelial dysplasia," *Journal of Biomedical Optics*, **11**(2), (2006).
41. Y. C. Wu and J. N. Y. Qu, "Autofluorescence spectroscopy of epithelial tissues," *Journal of Biomedical Optics*, **11**(5), - (2006).
42. M. Inaguma and K. Hashimoto, "Porphyrin-like fluorescence in oral cancer - In vivo fluorescence spectral characterization of lesions by use of a near-ultraviolet excited autofluorescence diagnosis system and separation of fluorescent extracts by capillary electrophoresis," *Cancer*, **86**(11), 2201-2211 (1999).
43. K. Konig, A. Ruck and H. Schneckenburger, "Fluorescence Detection and Photodynamic Activity of Endogenous Protoporphyrin in Human Skin," *Optical Engineering*, **31**(7), 1470-1474 (1992).
44. R. S. DaCosta, H. Andersson and B. C. Wilson, "Molecular fluorescence excitation-emission matrices relevant to tissue spectroscopy," *Photochemistry and Photobiology*, **78**(4), 384-392 (2003).
45. C. S. Betz, M. Mehlmann, K. Rick, H. Stepp, G. Grevers, R. Baumgartner and A. Leunig, "Autofluorescence imaging and spectroscopy of normal and malignant mucosa in patients with head and neck cancer," *Lasers in Surgery and Medicine*, **25**(4), 323-334 (1999).
46. H. M. Chen, C. Y. Wang, C. T. Chen, H. Yang, Y. S. Kuo, W. H. Lan, M. Y. P. Kuo and C. P. Chiang, "Auto-fluorescence spectra of oral submucous fibrosis," *Journal of Oral Pathology & Medicine*, **32**(6), 337-343 (2003).
47. D. C. G. de Veld, M. Skurichina, M. J. H. Witjes, R. P. W. Duin, D. J. C. M. Sterenborg, W. M. Star and J. L. N. Roodenburg, "Autofluorescence characteristics of healthy oral mucosa at different anatomical sites," *Lasers in Surgery and Medicine*, **32**(5), 367-376 (2003).
48. D. C. G. de Veld, M. Skurichina, M. J. H. Witjes, R. P. W. Duin, H. J. C. M. Sterenborg and J. L. N. Roodenburg, "Clinical study for classification of benign, dysplastic, and malignant oral lesions using autofluorescence spectroscopy," *Journal of Biomedical Optics*, **9**(5), 940-950 (2004).
49. D. C. G. de Veld, H. J. C. M. Sterenborg, J. L. N. Roodenburg and M. J. H. Witjes, "Effects of individual characteristics on healthy oral mucosa autofluorescence spectra," *Oral Oncology*, **40**(8), 815-823 (2004).
50. J. K. Dhingra, D. F. Perrault, K. McMillan, E. E. Rebeiz, S. Kabani, R. Manoharan, I. Itzkan, M. S. Feld and S. M. Shapshay, "Early diagnosis of upper aerodigestive tract cancer by autofluorescence," *Archives of Otolaryngology-Head & Neck Surgery*, **122**(11), 1181-1186 (1996).
51. A. Gillenwater, R. Jacob, R. Ganeshappa, B. Kemp, A. K. El-Naggar, J. L. Palmer, G. Clayman, M. F. Mitchell and R. Richards-Kortum, "Noninvasive diagnosis of oral neoplasia based on fluorescence spectroscopy and native tissue autofluorescence," *Archives of Otolaryngology-Head & Neck Surgery*, **124**(11), 1251-1258 (1998).

52. D. L. Heintzelman, U. Utzinger, H. Fuchs, A. Zuluaga, K. Gossage, A. M. Gillenwater, R. Jacob, B. Kemp and R. R. Richards-Kortum, "Optimal excitation wavelengths for in vivo detection of oral neoplasia using fluorescence spectroscopy," *Photochemistry and Photobiology*, **72**(1), 103-113 (2000).
53. D. R. Ingrams, J. K. Dhingra, K. Roy, D. F. Perrault, I. D. Bottrill, S. Kabani, E. E. Rebeiz, M. M. Pankratov, S. M. Shapshay, R. Manoharan, I. Itzkan and M. S. Feld, "Autofluorescence characteristics of oral mucosa," *Head and Neck-Journal for the Sciences and Specialties of the Head and Neck*, **19**(1), 27-32 (1997).
54. V. R. Kolli, H. E. Savage, T. J. Yao and S. P. Schantz, "Native Cellular Fluorescence Of Neoplastic Upper Aerodigestive Mucosa," *Archives of Otolaryngology-Head & Neck Surgery*, **121**(11), 1287-1292 (1995).
55. S. K. Majumder, N. Ghosh and P. K. Gupta, "Early diagnosis of oral cancer using autofluorescence spectroscopy and a probabilistic kernel machine," *Oral Oncology*, **1**(1), 100-100 (2005).
56. S. K. Majumder, N. Ghosh and P. K. Gupta, "Relevance vector machine for optical diagnosis of cancer," *Lasers in Surgery and Medicine*, **36**(4), 323-333 (2005).
57. S. K. Majumder, N. Ghosh and P. K. Gupta, "Support vector machine for optical diagnosis of cancer," *Journal of Biomedical Optics*, **10**(2), - (2005).
58. S. K. Majumder, N. Ghosh, S. Kataria and P. K. Gupta, "Nonlinear pattern recognition for laser-induced fluorescence diagnosis of cancer," *Lasers in Surgery and Medicine*, **33**(1), 48-56 (2003).
59. S. K. Majumder, A. Gupta, S. Gupta, N. Ghosh and P. K. Gupta, "Multi-class classification algorithm for optical diagnosis of oral cancer," *Journal of Photochemistry and Photobiology B-Biology*, **85**(2), 109-117 (2006).
60. S. K. Majumder, S. K. Mohanty, N. Ghosh, P. K. Gupta, D. K. Jain and F. Khan, "A pilot study on the use of autofluorescence spectroscopy for diagnosis of the cancer of human oral cavity," *Current Science*, **79**(8), 1089-1094 (2000).
61. R. J. Mallia, S. S. Thomas, A. Mathews, R. R. Kumar, P. Sebastian, J. Madhavan and N. Subhash, "Laser-induced autofluorescence spectral ratio reference standard for early discrimination of oral cancer," *Cancer*, (2008).
62. B. K. Manjunath, J. Kurein, L. Rao, C. M. Krishna, M. S. Chidananda, K. Venkatakrishna and V. B. Kartha, "Autofluorescence of oral tissue for optical pathology in oral malignancy," *Journal of Photochemistry and Photobiology B-Biology*, **73**(1-2), 49-58 (2004).
63. K. Onizawa, N. Okamura, H. Saginoya, H. Yusa, T. Yanagawa and H. Yoshida, "Analysis of fluorescence in oral squamous cell carcinoma," **38**(4), 343-348 (2002).
64. T. Tsai, H. M. Chen, C. Y. Wang, J. C. Tsai, C. T. Chen and C. P. Chiang, "In vivo autofluorescence Spectroscopy of oral premalignant and malignant lesions: Distortion of fluorescence intensity by submucous fibrosis," *Lasers in Surgery and Medicine*, **33**(1), 40-47 (2003).

65. R. A. Schwarz, W. Gao, D. Daye, M. D. Williams, R. Richards-Kortum and A. M. Gillenwater, "Autofluorescence and diffuse reflectance spectroscopy of oral epithelial tissue using a depth-sensitive fiber-optic probe," *Applied Optics*, **47**(6), 825-834 (2008).
66. D. C. G. de Veld, M. J. H. Witjes, H. J. C. M. Sterenborg and J. L. N. Roodenburg, "The status of in vivo autofluorescence spectroscopy and imaging for oral oncology," *Oral Oncology*, **41**(2), 117-131 (2005).
67. H. M. Chen, C. P. Chiang, C. You, T. C. Hsiao and C. Y. Wang, "Time-resolved autofluorescence spectroscopy for classifying normal and premalignant oral tissues," *Lasers in Surgery and Medicine*, **37**(1), 37-45 (2005).
68. M. Oheim, D. J. Michael, M. Geisbauer, D. Madsen and R. H. Chow, "Principles of two-photon excitation fluorescence microscopy and other nonlinear imaging approaches," *Advanced Drug Delivery Reviews*, **58**(7), 788-808 (2006).
69. P. Wilder-Smith, K. Osann, N. Hanna, N. El Abbadi, M. Brenner, D. Messadi and T. Krasieva, "In vivo multiphoton fluorescence imaging: A novel approach to oral malignancy," *Lasers in Surgery and Medicine*, **35**(2), 96-103 (2004).
70. S. M. Zhuo, J. X. Chen, X. S. Jiang, S. S. Xie, R. Chen, N. Cao, Q. L. Zou and S. Y. Xiong, "Layered-resolved microstructure and spectroscopy of mouse oral mucosa using multiphoton microscopy," *Physics in Medicine and Biology*, **52**(16), 4967-4980 (2007).
71. W. M. White, M. Rajadhyaksha, S. Gonzalez, R. L. Fabian and R. R. Anderson, "Noninvasive imaging of human oral mucosa in vivo by confocal reflectance microscopy," *Laryngoscope*, **109**(10), 1709-1717 (1999).
72. A. Ishimaru, *Wave propagation and scattering in random media*, Academic Press, New York, 1978
73. W. F. Cheong, S. A. Prahl and A. J. Welch, "A Review Of The Optical-Properties Of Biological Tissues," *IEEE Journal of Quantum Electronics*, **26**(12), 2166-2185 (1990).
74. T. J. Farrell, M. S. Patterson and B. Wilson, "A Diffusion-Theory Model Of Spatially Resolved, Steady-State Diffuse Reflectance For The Noninvasive Determination Of Tissue Optical-Properties Invivo," **19**(4), 879-888 (1992).
75. J. R. Mourant, J. P. Freyer, A. H. Hielscher, A. A. Eick, D. Shen and T. M. Johnson, "Mechanisms of light scattering from biological cells relevant to noninvasive optical-tissue diagnostics," *Applied Optics*, **37**(16), 3586-3593 (1998).
76. P. R. Bargo, S. A. Prahl, T. T. Goodell, R. A. Sleven, G. Koval, G. Blair and S. L. Jacques, "In vivo determination of optical properties of normal and tumor tissue with white light reflectance and an empirical light transport model during endoscopy," *Journal of Biomedical Optics*, **10**(3), - (2005).
77. A. Amelink, H. J. C. M. Sterenborg, M. P. L. Bard and S. A. Burgers, "In vivo measurement of the local optical properties of tissue by use of differential path-length spectroscopy," *Optics Letters*, **29**(10), 1087-1089 (2004).

78. D. Hidovic-Rowe and E. Claridge, "Modelling and validation of spectral reflectance for the colon," *Physics in Medicine and Biology*, **50**(6), 1071-1093 (2005).
79. I. S. Saidi, S. L. Jacques and F. K. Tittel, "Mie and Rayleigh Modeling of Visible-Light Scattering in Neonatal Skin," *Applied Optics*, **34**(31), 7410-7418 (1995).
80. R. L. P. van Veen, W. Verkruysse and H. J. C. M. Sterenborg, "Diffuse-reflectance spectroscopy from 500 to 1060 nm by correction for inhomogeneously distributed absorbers," *Optics Letters*, **27**(4), 246-248 (2002).
81. W. Verkruysse, G. W. Lucassen, J. F. deBoer, D. J. Smithies, J. S. Nelson and M. J. C. vanGemert, "Modelling light distributions of homogeneous versus discrete absorbers in light irradiated turbid media," *Physics in Medicine and Biology*, **42**(1), 51-65 (1997).
82. L. O. Svaasand, E. J. Fiskerstrand, G. Kopstad, L. T. Norvang, E. K. Svaasand, J. S. Nelson and M. W. Berns, "Therapeutic response during pulsed laser treatment of port-wine stains: Dependence on vessel diameter and depth in dermis," *Lasers in Medical Science*, **10**(4), 235-243 (1995).
83. C. H. Best, N. B. Taylor and J. B. West, *Best and Taylor's physiological basis of medical practice*, Williams & Wilkins, Baltimore, 1991
84. M. G. Muller, I. Georgakoudi, Q. G. Zhang, J. Wu and M. S. Feld, "Intrinsic fluorescence spectroscopy in turbid media: disentangling effects of scattering and absorption," *Applied Optics*, **40**(25), 4633-4646 (2001).
85. Q. G. Zhang, M. G. Muller, J. Wu and M. S. Feld, "Turbidity-free fluorescence spectroscopy of biological tissue," *Optics Letters*, **25**(19), 1451-1453 (2000).
86. J. Wu, M. S. Feld and R. P. Rava, "Analytical Model for Extracting Intrinsic Fluorescence in Turbid Media," *Applied Optics*, **32**(19), 3585-3595 (1993).
87. L. T. Perelman, V. Backman, M. Wallace, G. Zonios, R. Manoharan, A. Nusrat, S. Shields, M. Seiler, C. Lima, T. Hamano, I. Itzkan, J. Van Dam, J. M. Crawford and M. S. Feld, "Observation of periodic fine structure in reflectance from biological tissue: A new technique for measuring nuclear size distribution," *Physical Review Letters*, **80**(3), 627-630 (1998).
88. V. Backman, M. B. Wallace, L. T. Perelman, J. T. Arendt, R. Gurjar, M. G. Muller, Q. Zhang, G. Zonios, E. Kline, T. McGillican, S. Shapshay, T. Valdez, K. Badizadegan, J. M. Crawford, M. Fitzmaurice, S. Kabani, H. S. Levin, M. Seiler, R. R. Dasari, I. Itzkan, J. Van Dam and M. S. Feld, "Detection of preinvasive cancer cells," *Nature*, **406**(6791), 35-36 (2000).
89. C. Lau, O. Scepanovic, J. Mirkovic, S. McGee, C. C. Yu, S. Fulghum Jr., M. Wallace, J. Tunnell, K. Bechtel and M. Feld, "Vessel packaging should replace model-based light scattering spectroscopy in tissue spectroscopy," *Journal of Biomedical Optics*, (2008),(manuscript).

CHAPTER 4

Instrumentation and Data Analysis

In this chapter, we first present a detailed description of the instrumentation and software used to collect the spectroscopic data. Following this, we describe the methods used to calibrate and model the data in order to obtain quantitative parameters relating to the tissue morphology and biochemistry. Finally, we conclude with a description of the statistical methods that we employ in the analysis of the clinical data.

4.1 Instrumentation and Software

4.1.1 Basic Description of the FastEEM Clinical Instrument

Our laboratory has developed a portable, clinical device known as the Fast Excitation Emission Matrix (FastEEM)[1]. The FastEEM collects diffuse reflectance and fluorescence spectra in a fraction of a second under standard room illumination. A xenon arc flash lamp and a 308 nm XeCl excimer laser serve as the excitation light sources for the reflectance and fluorescence measurements, respectively. A photograph of the FastEEM in the clinical setting is shown in Figure 4.1.

Figure 4.1 A photograph of the FastEEM in the clinic.

Light is delivered to the tissue and collected via an optical fiber probe, which is placed in gentle contact with the tissue. The light returned from the tissue is focused onto the entrance slit of a diffraction spectrometer, dispersed onto an intensified CCD, and stored by a computer. The operation of the instrument and data acquisition is controlled by a custom-designed software program.

4.1.2 Generation of Excitation Light

A schematic diagram showing the major hardware components of the FastEEM is shown in Figure 4.2.

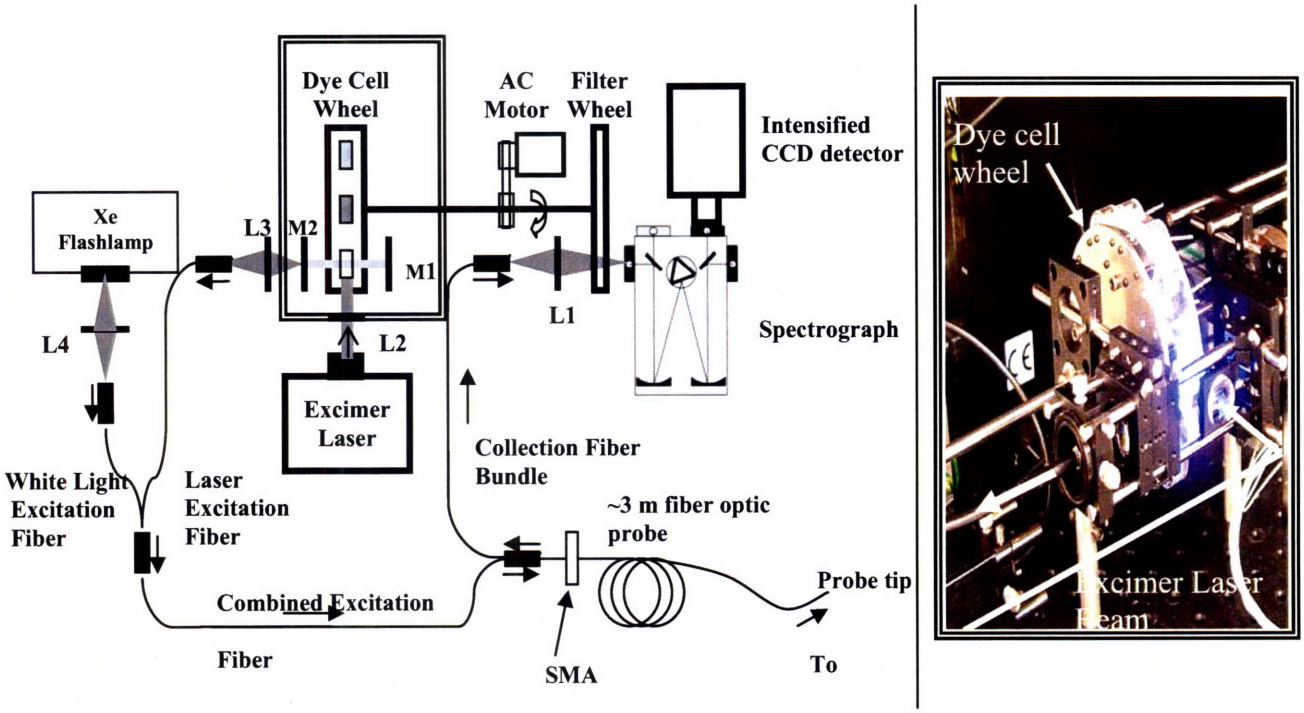

Figure 4.2 Diagram showing the major hardware components of the FastEEM. The Xe flashlamp and XeCl excimer laser are used to generate the excitation light for the reflectance and fluorescence measurements, respectively. The excitation light is coupled into the delivery fiber of the optical fiber probe. The returned light is directed through the filter wheel, to the spectrograph, and CCD. A photograph of the dye cell wheel shown in the schematic on the left (boxed area) is shown on the right.

The flashlamp (60 W, L7684, Hamamatsu, Bridgewater, NJ) generates broadband white light over the wavelength range of approximately 300-800 nm. The XeCl excimer laser (OPTEX, Lambda Physik, Ft. Lauderdale, FL), provides fluorescence excitation light at 308 nm and nine additional wavelengths, by pumping a series of dyes with emission wavelengths ranging from 340-460 nm. Table 4.1 lists the nine dyes used to generate the fluorescence excitation light, their emission wavelengths, and the full width at half maximum (FWHM) of the signal. The dyes are

contained within cylindrical cuvettes (d = 22 mm, path length = 12.5 mm) situated on a rotating wheel (dye wheel). The dye wheel rotates at ~3 Hz. The cuvettes are made of optically transparent Spectrosil Quartz (Starna Cells, Inc., Atascadero, CA). In order to generate each excitation wavelength, the laser beam is directed by a series of mirrors and focused to a point, such that as the dye wheel rotates, each dye is successively brought into the path of the beam and excited during the collection period. For 308 nm fluorescence excitation light, a cylindrical lens is positioned on the dye wheel (no cuvette is present) and the output from the excimer laser is conveyed directly to the delivery fiber of the optical fiber probe. The FWHM of the 308 nm excitation light is 8.5 nm.

Laser Dye	Solvent	Emission λ [nm]	FWHM [nm]
p-Terphenyl (PTP)	*p*-dioxane	341.6	5
PBD	*p*-dioxane	360.0	5
Exalite 384	*p*-dioxane	383.6	6
Exalite 398	*p*-dioxane	400.0	6
LD 425	*p*-dioxane	407.2	12
Stilbene 420	methanol	425.7	8
Coumarin 440	methanol	441.0	10
Coumarin 460	methanol	461.5	10
Coumarin 480	methanol	483.1	9

Table 4.1 Properties of the dyes used to generate the nine fluorescence excitation wavelengths.

Emission light generated from the reflectance and fluorescence sources, respectively, are focused into two separate ends of a Y-fiber coupler. The fibers from the two excitation sources are merged and form the central delivery fiber of the optical fiber probe. The flashlamp and excimer laser are both pulsed sources. The pulse duration of the flashlamp and excimer laser is 2.9 μs and 8 ns, respectively, at FWHM. The collection of a complete data set entails the sequential excitation and

collection of the reflectance and fluorescence data and takes ~0.3 s. Figure 4.2 summarizes the specifications for the two light sources.

Property	Flash Lamp	Excimer Laser
Repetition Rate (s^{-1})	6	30
Radiant Energy ($pulse^{-1}$)	1 μJ	1 -10 μJ
Pulse Width (FWHM)	2.9 μs	8 ns
Irradiance per Pulse (kW/cm^2)	0.03	184-1,840
Fluence per Pulse (mJ/cm^2)	0.17	0.17-1.7

Table 4.2 Summary of the specifications for the two light sources preent on the FastEEM, the Xe arc flashlamp and XeCl excimer laser.

4.1.3 Optical Fiber Probe

An optical fiber probe is used to deliver the excitation light from the two sources, as well as to convey light emitted from the tissue to the detection components. The probe consists of a central delivery fiber, into which the excitation light is coupled, surrounded by 6 collection fibers in a closest-packed configuration. The fiber core is composed of fused silica and has a diameter of 200 μm (Numerical Aperture (NA) = 0.22). With the cladding, the diameter of each fiber is 220 μm. The entire 3 m length of the probe is enclosed in Teflon tubing for protection. At the probe tip, the fibers are exposed

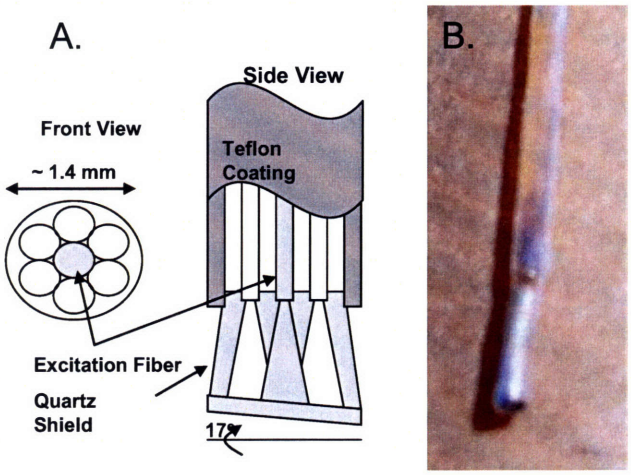

Figure 4.3 (a) Diagram of the optical fiber probe showing a front and side view (b) Photograph showing the tip of the optical fiber probe, which is placed in contact with the tissue.

from the cladding, fused, embedded in a protective silver epoxy, and surrounded by a protective sheath. A 1.5 mm long quartz shield is placed at the very tip to create a reproducible geometry when in contact with the tissue. With the cladding and protective sheath, the probe tip has a diameter of approximately 1.4 mm. The tip of the probe is polished to an angle of approximately 17° to prevent specular reflections from entering the probe and producing artifacts in the measured signal. Figure 4.3 shows a schematic of the probe geometry [Fig 4.3a] and a photograph of the probe [Fig 4.3b]. The diameter of the delivered excitation light is approximately 480 μm.

At the distal end of the probe, the 6 collection fibers separate from the source fiber and couple to the detection components via a set of 6 SubMiniature version A (SMA) connectors. Similarly, the distal end of the delivery fiber is coupled to the excitation sources via a separate SMA connector.

4.1.4 Detection System

The six collection fibers are aligned vertically, focused, and directed through a filter on a second rotating wheel (filter wheel). The filter wheel and dye wheel are coupled so that they rotate in parallel and the appropriate long pass filter for the particular excitation light in the sequence is in place. The filter wheel contains 10 filters which specifically block back-reflected excitation light for each fluorescence excitation wavelength.

After passing through the filter, the collected light is directed through the slit of the spectrograph (Spectra Pro 150 Acton Research, Acton, MA, 300 grooves/mm with 500 nm blaze, center wavelength 550 nm), where it is dispersed and then

detected by an intensified CCD (PIMAX 1024RB Roper Scientific, Princeton, NJ). The dimensions of the chip are 256x1024 pixels. The rapid gated detection of the intensified CCD minimizes motion artifacts and enables data to be acquired in the presence of ambient light. The CCD is cooled to −20° C by Peltier (thermoelectric) cooling to reduce noise from the dark current (thermally generated charge) and improve the signal-to-noise ratio. The wavelength resolution of the system based on the FWHM is approximately 5.3 nm.

4.1.5 Timing and Triggering

The timing of the data acquisition is controlled by a National Instruments timer/counter board (PCI-6602, National Instruments, Austin, TX). A ring of metal pins attached to the dye wheel are positioned to sequentially interrupt a beam of light between two optical sensors (Omron Photomicrosensors EE-SX770, Schaumburg, IL) as each dye (or blank position) is rotated into the path of the laser. When a pin passes through the beam, the constant voltage signal generated between the two sensors decays, generating a trigger signal that is read by the timer board. The timer board then sends an appropriate trigger signal to the flash lamp or the laser. A second ring of pins on the dye wheel controls the gating of the CCD via a second set of optical sensors.

4.1.6 Software

A custom-designed software program created using LabView 7.0 (National Instruments Corporation, Austin, TX) is used to control the hardware components and coordinate the timing of the initiation and termination of data collection. The software was developed by engineers in our laboratory. The program is operated through a graphical user (GUI) interface on the FastEEM's computer. Using the GUI, various settings can be controlled including the gain on the CCD and the number of spectra to be averaged for each measurement. Real-time analysis of the collected diffuse reflectance and fluorescence data using the mathematical models described in Chapter 3 can also be performed by the software in combination with MATLAB (The MathWorks, Inc. Natick, MA).

A complete set of data consists of two diffuse reflectance spectra and an excitation-emission matrix (EEM), which contains fluorescence intensities as a function of the excitation and emission wavelengths. As described earlier, a single data set can be collected in ~0.3 s. Because we typically average 5 sets of reflectance and fluorescence data for each saved measurement, the collection time is on the order of ~1.5 s. For each measurement, the standard deviation of the 5 spectra as a function of the emission wavelength is also recorded.

4.1.7 Advances in Instrumentation

Reliable Measurement of Diffuse Reflectance

At the beginning of this project, we worked extensively to identify the source of an artifact present in the reflectance data. Unlike earlier data, recent data began to show a large deviation in the red region, starting at approximately 600 nm (see Fig. 4.4(a)), between the modeled reflectance spectrum and *in vivo* data. The data measured by the FastEEM systematically indicated higher amounts of reflected light than the modeled data in this wavelength regime. In order to elucidate the source of this artifact, we carried out studies with physical tissue models (solutions of polystyrene beads and hemoglobin in water with known scattering and absorption parameters), so that the actual data could be compared with the theoretical modeled reflectance spectrum.

Early steps in the investigation of the source of this problem involved testing several hypotheses to explain the extra counts including the following: 1) Limitations in the resolution of the spectrophotometer resulted in broadening of peaks and the distribution of extra counts into a wide range of wavelengths, and 2) The additional counts were due to second order light generated from the diffraction grating.

To test the first hypothesis, we took a background spectrum in which we shone the probe light into a black cup of water. We then directed light from a laser pointer (630-680 nm) into the probe and examined the collected light. If our first hypothesis were true we would expect to see an increase in the background counts on either side of the laser line due to scattering of this light into other wavelengths,

compared to a spectrum collected when the laser light was absent. The results showed no difference in the backgrounds.

Second-order light corresponds to a higher order diffraction pattern of the signal from the shorter wavelengths, produced at exactly twice the original wavelength. Because of the spatial overlap of this signal with the true reflectance measured at the longer wavelengths, the counts recorded by the CCD would represent a combination of the two signals. To test the possibility that second order light from the grating was contributing to the observed signal, we collected spectra from phantoms after placing a 335 nm, 365 nm or, no long-pass filter in front of the flash lamp. We expected less deviation at the higher wavelengths with the filters present because they would cut off the transmission of the first order light, and furthermore, the most improvement would be seen in the case of the 365nm long-pass filter. Figure 4.4 shows the results from this experiment. The abnormal rise in the reflectance at the longer wavelengths is clearly apparent in Figure 4.4(b). With the 335 nm long pass filter the rise does not occur until ~650 nm which is approximately twice the wavelength at which this filter begins to transmit light. Similarly, the 365 nm long pass filter further suppresses the rise in the reflectance until ~720 nm.

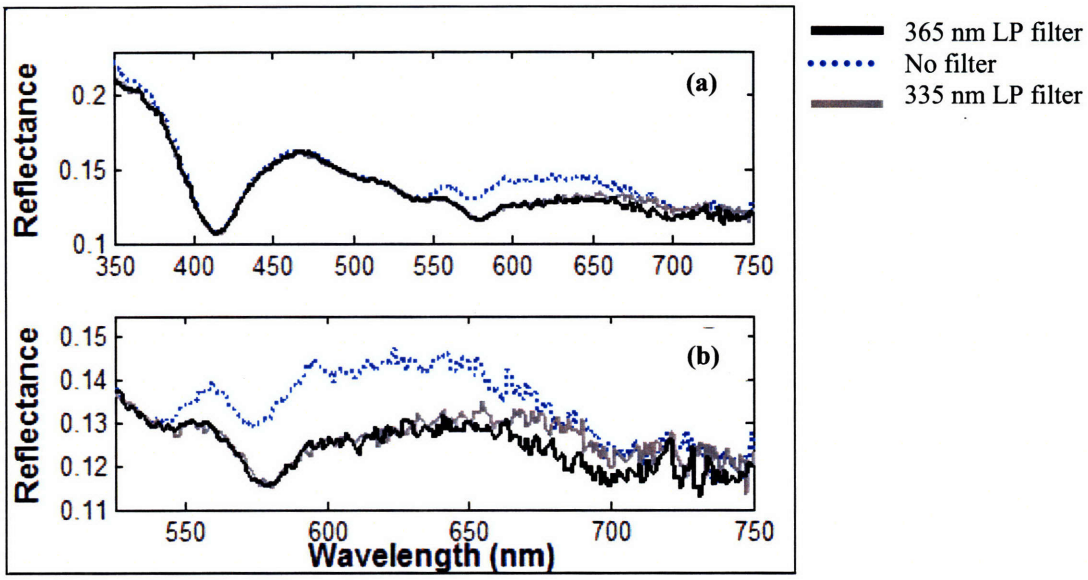

Figure 4.4 (a) Calibrated reflectance for a tissue phantom taken with no filter present (blue dotted line), with a 335 nm long-pass filter (grey line), or with a 365 nm long-pass filter (black line) placed in front of the flash lamp. (b) A close-up of the reflectance spectra from (a) at longer wavelengths.

This experiment confirmed that the problem was due to second order light. Our system is particularly affected by this artifact because data is collected over a broad wavelength range (300 – 800 nm).

After isolating the source of the problem, we set out to characterize the magnitude of the second order effect. An experiment was carried out in which a series of fluorescent solutions were excited by the laser, thus providing well-defined fluorescence emission peaks. The peak emission wavelengths of the dyes ranged from approximately 350nm to 380 nm. We were able to observe a second well-defined peak with a smaller magnitude at longer wavelengths with a center wavelength at approximately twice the center wavelength of the sample's fluorescence peak as shown in Figure 4.5.

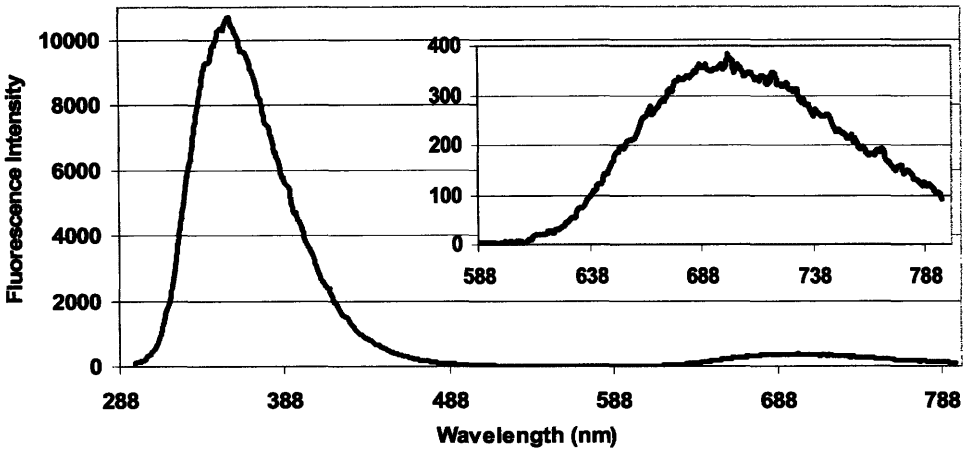

Figure 4.5 Plot of the measured signal from a fluorescent dye with an emission peak at ~360nm. The second order light results in a second peak at approximately twice the peak wavelength. The inset shows a zoomed in view of the fluorescence signal detected at longer wavelengths.

By aligning the signal of the second order light on the scale of the first order light, [Figure 4.6(a)], and dividing the second order light by the first order light, we developed a wavelength-dependent correction curve. By combining the ratios obtained from the various fluorescence emission peaks, the correction factors for a wide range of wavelengths could be fit by a linear equation ($R^2 = 0.97$) as shown in Figure 4.6(b).

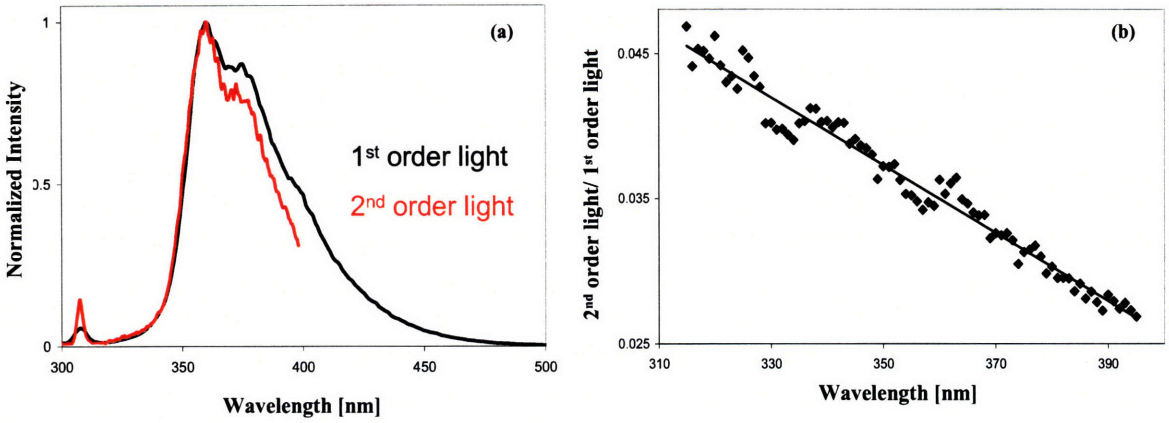

Figure 4.6 (a) Plot of the first order light and the second order light (plotted at half the actual wavelength at which it was detected). Both signals were normalized to the peak. (b) Plot of the correction factor, the ratio of the second order light to the first order light. A linear fit is also shown.

The introduction of the second order light in the reflectance signal was caused by a switch to UV grade optics, which allowed shorter wavelength light from the flash lamp (down to ~160 nm) to be transmitted through the system. Since the flash lamp produces light as far down as ~270 nm, second order light was generated beginning at ~540 nm.

Our next experiments tested whether the linear correction function ("second order equation") could be used to fix previously collected data. We compared phantoms taken without the filter in front of the flashlamp and fixed using the second-order equation, to the signal measured from the same phantom after placing a 400 nm long pass filter in front of the flash lamp. This filter would block all second order light generated below ~760 nm. Figure 4.7 shows the results for one phantom, the results of which confirmed that the two methods are equivalent.

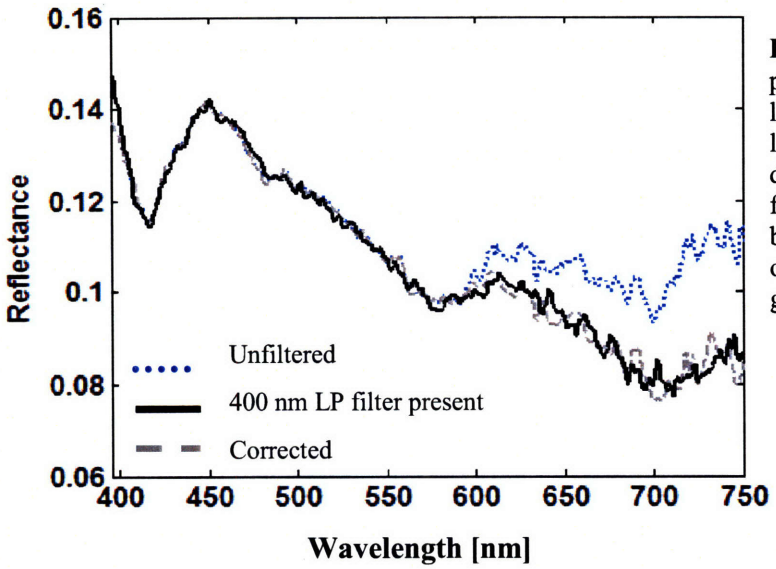

Figure 4.7 Reflectance from a phantom taken with a 400 nm long pass filter (black solid line), without the filter (blue dotted line), and without the filter, and without the filter, but corrected by the second order fix equation (dashed grey line).

As a permanent solution, we created a hardware fix for the problem. In a single data acquisition, the reflectance measurement is performed twice in order to correct for second-order light from the spectrograph that contributes to the diffuse reflectance signal measured by the detector.

In the first reflectance measurement, the filter wheel slot is empty, therefore the entire measured reflectance signal enters the spectrograph. In this case, the signal from shorter wavelengths is reliable, but at longer wavelengths (>550 nm), the second-order light generated by the grating of the spectrograph also contributes to the signal measured by the CCD. For the second measurement, the filter wheel contains a 435 nm long-pass filter, which prevents transmission of the signal at shorter wavelength from being transmitted to the spectrograph. As a result, the measured reflectance at the longer wavelengths represents the true measured signal from the tissue. In the preprocessing of the collected reflectance, the data from the 2 reflectance measurements are merged. The merging process involves first correcting the second (filtered) measurement for the transmission properties of the filter, and then scaling it to overlap with the measurement from the first (unfiltered)

measurement. The final spectrum contains the signal measured from the unfiltered reflectance data at shorter wavelengths and the counts from the filtered reflectance data at longer wavelengths.

4.2 Data Preprocessing and Analysis

4.2.1 Artifact in the Measured Data

In the course of analyzing the clinical data, we observed a signal in the fluorescence data which was not consistent with the fluorescence of known fluorophores, and also appeared in the fluorescence at every excitation without any shift in wavelength. It was discovered that the white light emitting source worn by the physician ("headlight") and the overhead lights in the operating room which were shone directly on the tissue during data collection, were diffusing into the tissue and being collected by our probe. Unlike ambient room light which is weak and not focused directly on the tissue, these light sources are directed on the tissue in a more focused manner. We devised a means of correcting both the reflectance and fluorescence signal for this artifact as follows. First we identified a fluorescence excitation at which the measured fluorescence excitation signal was negligible or weak (often 425 or 440 nm excitation). Because the measured signal for this excitation would not be due to fluorescence, it provided a means of isolating the stray light signal. By correcting this stray light spectrum for the transmission properties of the long-pass filter used for this specific excitation light, we obtained the true

spectrum of the stray light. The reflectance spectra were corrected by subtracting these counts from the measured spectrum. The measured signals at the remaining fluorescence excitations were then corrected by subtracting this spectrum from the measured signal at the excitation after multiplying the stray light spectrum by the transmission properties of the filter used at the particular excitation, since the measured signal must pass through the filter (on the filter wheel, Section 4.1.4)) before it is detected.

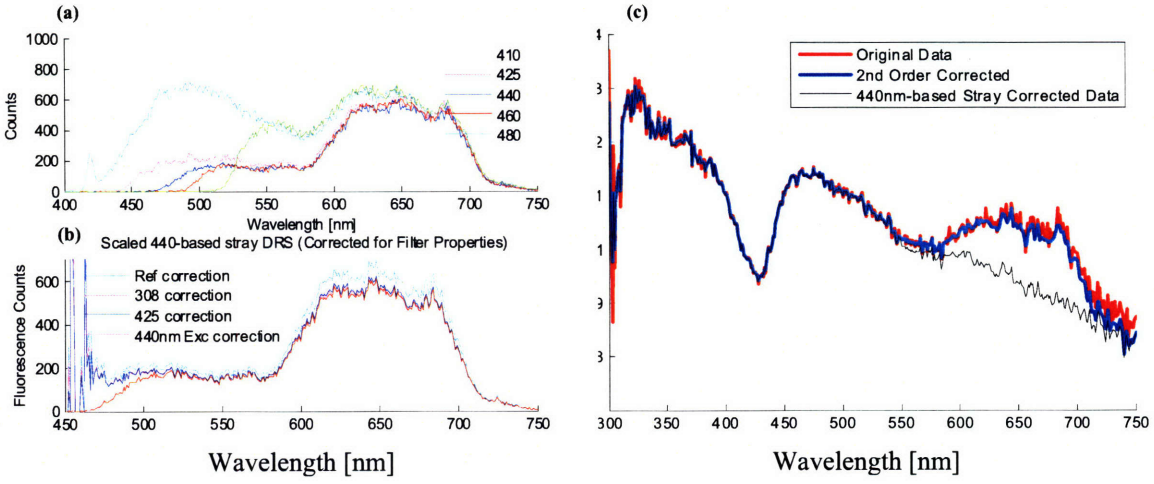

Figure 4.8 (a) Measured fluorescence signal for 5 fluorescence excitations, which contains the stray light (b) Examples of the stray light signal isolated from the reflectance and 3 fluorescence excitations after correcting for the transmission properties of the filters (fluorescence only) (c) application of second order light correction (merging of unfiltered and filtered spectra) (blue line), and stray light correction (black line), to calibrated reflectance (red line).

Figure 4.8(a) shows the measured fluorescence counts after correcting for the background signal for 5 fluorescence excitations, with clearly visible signs of stray light. Figure 4.8(b) shows the isolated stray light signal to correct the measured reflectance, as well as the correction for 3 fluorescence excitations. Because the stray light signal isolated from each of the fluorescence excitations has been corrected for

the transmission properties of the respective filters, the spectra all overlap. Figure 4.8(c) shows the uncorrected calibrated (original) reflectance data, and the effect of the stray light correction. In addition the second order correction was performed by merging the data. The aforementioned correction was used to correct data which had already been collected. Once this was recognized, extra light sources other than room lights were turned off prior to each measurement. For a large number of samples, the fluorescence spectra excited at the longer wavelengths were imperfectly corrected (overcorrection was frequently noted), likely due to the fact that the fluorescence emission and stray light overlap to a greater extent, and the presence of absorption features in the stray light, which would remove rather than add light to the measured signal. This meant that the measured signal for most of the 10 fluorescence excitations were unreliable. As a result, only the 308 nm and 340 nm excitations were analyzed in the work described in this thesis as their emission was reliable.

4.2.2 Calibration and Modeling of the Reflectance Data

All data were corrected for background reflectance and fluorescence by measuring a spectrum from water in a black container and subtracting these counts from the signal measured from the tissue and calibration spectra. The calibrated diffuse reflectance spectrum, $R_{cal}(\lambda)$, was generated by dividing the measured spectrum, $R_{meas}(\lambda)$, by a spectrum measured from a Spectralon (LabSphere, Inc., North Sutton, NH) 20% reflectance standard, $R_{std}(\lambda)$, as follows:

$$R_{cal}(\lambda) = S_R(\lambda) \cdot [R_{meas}(\lambda) - R_{bg}(\lambda)] / [R_{std}(\lambda) - R_{bg}(\lambda)]. \quad (4.1)$$

$S_R(\lambda)$ denotes the fraction of the light reflected by the spectralon reflectance standard and is an intrinsic property of the material. The values for $S_R(\lambda)$ were provided by the manufacturer. R_{bg} refers to the background reflectance counts. As shown in the equation, all of the quantities are wavelength dependent. Calibration of the measured reflectance corrects for the spectral profile of the flashlamp and the wavelength-dependence of the detector response.

A non-linear least-squares fitting algorithm was used to fit the calibrated reflectance data to the model of Zonios *et al.* described in the previous chapter. The function minimizes X^2, the error between the modeled fit, R_{fit}, and R_{cal}, where

$$\chi^2 = \sum[(R_{fit}(\lambda) - R_{cal}(\lambda))^2 / \sigma_{R_{cal}}(\lambda)^2]. \quad (4.2)$$

The standard deviation of R_{cal}, σ_{Rcal}, was calculated by propagating the error (standard deviation) in each of the measured spectra used to calculate the calibrated data according to the following equation:

$$\sigma_{R_{cal}}(\lambda) = S_R \cdot \left[\left(\frac{\partial R_{cal}}{\partial R_{meas}} \right)^2 \sigma_{R_{meas}}{}^2 + \left(\frac{\partial R_{cal}}{\partial R_{bg}} \right)^2 \sigma_{R_{bg}}{}^2 + \left(\frac{\partial R_{cal}}{\partial R_{std}} \right)^2 \sigma_{R_{std}}{}^2 \right]^{1/2} \quad (4.3)$$

The standard deviations are denoted by σ. All values of σ are in units of counts except for σ_{Rcal}, which is in calibrated reflectance units The standard deviation for

each measurement is derived from the 5 individual spectra which are averaged to produce a single recorded measurement. Based on the previous equation, the error in the calibrated reflectance (in reflectance units) was computed as follows:

$$\sigma_{R_{cal}}(\lambda) = \frac{S_R}{(R_{std} - R_{bg})^2} \cdot \left[(R_{std} - R_{bg})^2 \sigma_{R_{meas}}{}^2 + (R_{meas} - R_{std})^2 \sigma_{R_{bg}}{}^2 + (R_{bg} - R_{meas})^2 \sigma_{R_{std}}{}^2 \right]^{1/2}. \quad (4.4)$$

4.2.3 Calibration and Modeling of the Fluorescence Data

The fluorescence emission measured by the system at a specific excitation wavelength, $F_{meas}(\lambda_x,\lambda)$, is a function of the laser intensity, I_x, detector response, S_{system}, and the intensity of the fluorescence emitted by the fluorophore, F_{fl}:

$$F_{meas}(\lambda_x, \lambda) = I_x(\lambda) \cdot F_{fl}(\lambda_x, \lambda) \cdot S_{system}(\lambda). \quad (4.5)$$

The dependence of these quantities on the excitation wavelength and/or emission wavelength is indicated by λ_x and λ, respectively. Because we do not measure the intensity of the laser during data collection and variations from day to day can affect the measured fluorescence, we calibrate the fluorescence intensity using a standard solution of rhodamine 610 dye (Exciton Dayton, OH) in ethylene glycol. Rhodamine has a single emission peak at 580 nm and can be excited by all the excitation wavelengths. The goal of the rhodamine calibration procedure is to isolate the intensity of the fluorescence emitted by the tissue from the intensity of the excitation light generated from the laser and dyes, as well as other variations in the system. This

calibration also serves as a means of relating the intensity of the excitation light for the different excitation wavelengths.

By measuring the fluorescence emission profile of the rhodamine solution using a fluorimeter, we determined the true relative fluorescence intensities for each excitation wavelength. We then normalized the intensity values measured at each excitation wavelength by the intensity measured at 308 nm excitation. Each day of data collection, we measure the same rhodamine solution on the FastEEM, determine the maximum peak intensity for each excitation wavelength, and then apply the same normalization process. Using the normalized intensity ratios obtained from the fluorimeter, $S^{fl}_{reference}$, and the set of normalized ratios derived from the measurement with the FastEEM, S^{fl}_{rhod}, we apply a correction factor to the measured fluorescence and obtain the calibrated fluorescence. For each excitation wavelength, the calibrated fluorescence, F_{cal}, is calculated according to the following equation:

$$F_{cal}(\lambda_x, \lambda) = [F_{meas}(\lambda_x, \lambda) - F_{bg}(\lambda_x, \lambda)] \cdot [(S^{fl}_{reference}(\lambda_x) / S^{fl}_{rhod}(\lambda_x)] . \quad (4.6)$$

The background fluorescence is denoted by F_{bg}. Although the rhodamine fluorescence standard corrects the fluorescence intensity and enables the comparison of the intensity across different excitation wavelengths, despite different output energies from the dyes, it does not correct for distortions in the measured fluorescence lineshape.

The extraction of the intrinsic fluorescence is achieved using the IFS algorithm presented in Chapter 3.3. The spectra of the component fluorophores (basis

spectra) for each excitation wavelength were then extracted from the intrinsic fluorescence using multivariate curve resolution (MCR) and in some cases, a spectrum collected from the pure chemical. The intrinsic fluorescence spectrum extracted for each excitation wavelength was modeled as a linear combination of the component fluorophores using a linear least squares fitting algorithm in order to obtain the contribution of each fluorophore. The component fluorophores for each excitation wavelength are shown in Table 4.3, where an "x" denotes the component is included in the model.

Excitation Wavelength [nm]	Tryptophan	Collagen	NADH
308	x	x	x
340		x	x

Table 4.3 Summary of the fluorophores excited at each excitation wavelength analyzed in the clinical studies.

4.2.4 Wavelength Calibration

An emission spectrum from a mercury argon lamp (HG-1, Ocean Optics, Dunedin, FL) was recorded each day of data collection. Because the wavelengths of these peaks are known, we can determine the relationship between pixel number and wavelength by fitting a second-order polynomial to a plot of the wavelength of each peak versus the pixel number of the peak.

4.3 Statistical Methods

4.3.1 Logistic Regression

Logistic regression is a classification method in which one or more independent predictor variables (spectral parameters, in our case) are used to predict the probability (dependent variable), of a sample's membership to one of two classes (binary outcome)[2]. The probability, p, as a function of the independent variables, $X_1 \ldots X_n$, is modeled according to the following equation:

$$\ln\left(\frac{p}{1-p}\right) = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \ldots \beta_n X_n, \tag{4.7}$$

where α is a constant, and $\beta_1 \ldots \beta_n$ represent fit coefficients for each independent variable. The values of these quantities are obtained by applying the method of maximum likelihood estimation (MLE) on a set of training data in which the class membership is known. MLE maximizes the likelihood function, the joint probability or likelihood of obtaining the observed data, in order to extract α and the coefficients for each of the parameters[3]. The MLE function is solved iteratively by a computer. In order to obtain p, this equation can be rearranged as follows:

$$p = \frac{1}{1 + e^{-(\alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \ldots \beta_n X_n)}} \tag{4.8}$$

The value of p indicates the probability of a positive (e.g. diseased) outcome. A more in depth description of logistic regression can be found in the following source[3].

4.3.2 Identification of Significant/Diagnostic Parameters

While we extract a number of spectral parameters from the reflectance and fluorescence spectra, not all of these parameters contribute significantly to the separations of interest or provide diagnostic information. In order to identify the most relevant parameters to be used for constructing the spectral models, we employ the likelihood ratio test (also known as the log likelihood ratio test). The test is based on a comparison of a measure of the "badness of the fit" for the logistic regression models developed from two different sets of parameters. This measure is known as the deviance, D, and is related to the maximum likelihood value of a fitted model, L_i, as follows:

$$D = -2 \cdot \ln L_i \tag{4.9}$$

The deviance is an output obtained from the numerical calculation performed by the computer to maximize the likelihood function. The log likelihood ratio test (LLR) is calculated as the difference in the deviances for two models, or equivalently, using the following ratio:

$$LLR = -2 \cdot \ln\left(\frac{L_1}{L_2}\right) \tag{4.10}$$

L_1 represents the maximum likelihood value for the reduced or simpler model, and L_2 the value for the full model. For this test, the two models L_1 and L_2 must be nested, meaning that the parameters in the reduced model are a subset of those in the full model.

The LLR tests the null hypothesis that the parameter omitted in the simple model has a coefficient of 0 in the logistic regression model. The higher the value of the LLR, the more significant the omitted parameter. A non-significant likelihood ratio test indicates no difference between the full and reduced models; therefore, the specific regression coefficient(s) being tested can be omitted. The LLR approximately follows a chi-squared distribution with degrees of freedom equal to the difference in the number of parameters present in the full model and reduced model. Based on the values of the LLR, all the parameters can be ranked in terms of their importance in contributing to the model; however, only those parameters for which the test statistic exceeds the critical chi-squared value at the 5% significance level are considered statistically significant.

4.3.3 Leave-one-out Cross Validation and Receiver Operator Characteristic Curves

In order to assess the effectiveness with which various diagnostic parameter combinations and algorithms separate two classes (discrimination), we used leave-one-out cross-validation (LCV). LCV is a method applied in cases in which the number of samples available for constructing a training set and test set for the evaluation of the performance of a classifier is limited. LCV is performed by partitioning the total number of samples, n, into two groups: a single test case, and a training set comprised of the remaining n-1 samples. The diagnostic algorithm developed from the training set is then applied to the test case, and this procedure is then repeated so that each of the n samples are sequentially omitted and treated as the test case.

One set of measures of the discriminatory power or accuracy of a diagnostic algorithm is the sensitivity and sensitivity[4]. The sensitivity indicates the percentage of correctly classified positive (diseased) cases. Sensitivity does not take into account false positive cases. In contrast, specificity represents the percentage of correctly predicted negative (non-diseased) cases. By summing the number of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) results for the classification of the final point across all the n samples, the sensitivity and specificity can be calculated according to the following equations:

$$\text{Sensitivity} = TP / (TP + FN), \tag{4.11}$$

$$\text{Specificity} = TN / (TN + FP). \tag{4.12}$$

TP corresponds to the number of positive (diseased) samples that were correctly classified by the diagnostic algorithm as positive, while FN is the number of times in which a positive sample was incorrectly classified as negative (non-diseased). TN is the number of negative samples correctly classified by the diagnostic test as negative, while FP is the number of negative samples which were incorrectly classified as positive by the test. A perfect separation, as could be achieved for two non-overlapping groups, would have a sensitivity and specificity of 100% and 100%, respectively.

By computing the sensitivity and specificity based on the LCV results for each value of the p (probability threshold), we can generate a receiver operator characteristic (ROC) curve in which the sensitivity (true positive rate) is plotted against 1-specificity (false positive rate)[4]. This curve depicts the trade off between the sensitivity and specificity for various thresholds in p. Each point along the ROC curve

corresponds to the sensitivity and specificity for a specific threshold. The optimal sensitivity and specificity, or combination, for the specific application is chosen and used to report the accuracy of the diagnostic algorithm. The final decision line/surface for diagnosis would then represent the solution to the logistic regression equation (determination of the coefficients of the independent variables) calculated at the specific p value corresponding to this point. A new test case which produces a p greater than this value would be considered a positive test result.

In addition to the sensitivity and specificity, a diagnostic test can also be evaluated in terms of the positive predictive value (PPV) and negative predictive value (NPV). The PPV represents the probability of disease among people who test positive, and the NPV, the probability of the absence of disease among people who test negative. They can be calculated according to the following equations:

$$PPV = \frac{sensitivity \cdot prevalence}{sensitivity \cdot prevalence + (1 - specificity) \cdot (1 - prevalence)} \tag{4.13}$$

$$NPV = \frac{specificity \cdot (1 - prevalence)}{(1 - sensitivity) \cdot prevalence + specificity \cdot (1 - prevalence)} \tag{4.14}$$

The PPV and NPV provide a measure of the usefulness of the test given the prior probability of disease in the population before the test was administered[5].

Using the ROC curve, another measure of the accuracy of a diagnostic test can be extracted by calculating the area under the curve (ROC-AUC)[6]. The ROC-AUC corresponds to the probability that a positive sample would exhibit a higher classifier

score than that of a negative sample, if a random pair of samples is selected from the positive and negative groups[4]. A value of the ROC-AUC equal to 0.5 indicates that the test produces diagnostic results equivalent to what would be observed due to chance alone. At the other extreme, a value of 1.0 indicates no overlap in the distributions and thus a perfect separation. The higher the ROC-AUC value, the better the diagnostic accuracy. This measure characterizes the overall accuracy of the model across the entire range of thresholds, and thus is more comprehensive than the sensitivity and specificity, which are determined from a single point. The quality of the ROC curve, and thus the ROC-AUC, increases with the number of thresholds tested (step size) and with increasing size of the test sets[4]. An example of an ROC curve is shown in Figure 4.9. Curves depicting an ROC-AUC of 0.5 and 1.0, respectively, as well as a typical ROC curve for a diagnostic test are shown.

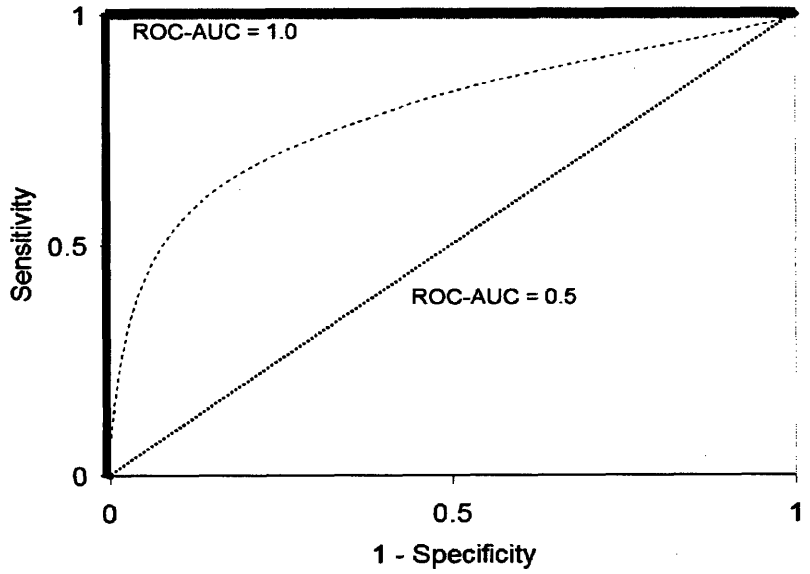

Figure 4.9 Examples of ROC curves. The dotted line represents the ROC curve that would result for a test which produces a classification result equivalent to what would be observed due to chance (ROC-AUC = 0.5). The thick solid line shows the ROC curve for a test which produces a perfect separation (ROC-AUC = 1.0). The dashed line shows a typical curve for a classification test with a ROC-AUC of >0.70.

4.3.3 *Interquartile Range (IQR) Exclusion Criteria*

On method of identifying outliers in data is based on the interquartile range (IQR) criteria. In order to understand the IQR exclusion criteria, it is helpful to introduce the box plot. A basic box plot summarizes the following properties of a set of data: the median, the lower limit (minimum), upper limit (maximum), the lower quartile (median of the lower half of the data), and upper quartile (median of the upper half of the data)[7]. In some box plots, notches are also displayed, which represent the approximate confidence interval about the median. The IQR is defined as the difference between the value of the upper quartile (Q3) and the lower quartile (Q1). The IQR represents the middle 50% of the data points, since a quartile corresponds to the range within which 25% of the data points lie. The length of a box plot is equivalent to the IQR. Figure 4.10 shows an example of a box plot with the major components labeled.

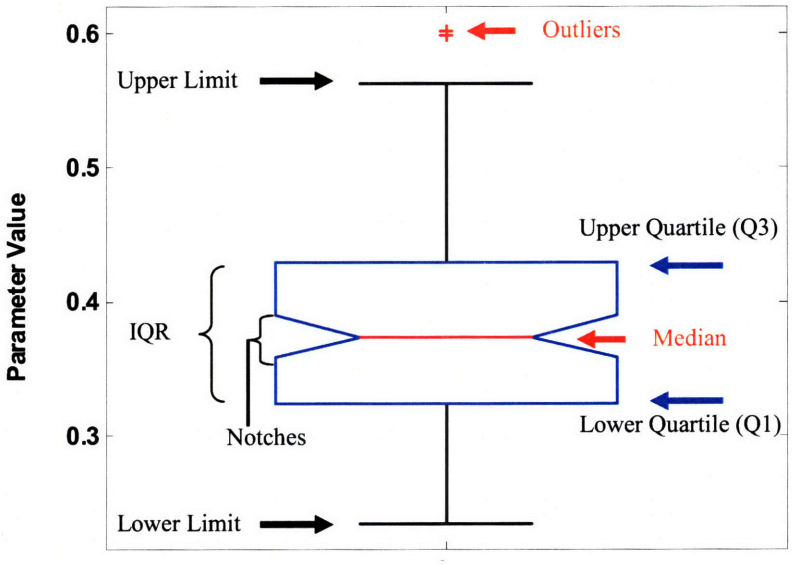

Figure 4.10 Example of a box plot. The major components of this plot are labeled, including the median, upper and lower limits, upper and lower quartiles, notches, outliers, and IQR.

According to the IQR exclusion criteria, data points which exceed Q3 by 1.5·IQR, or fall below Q1 by 1.5·IQR are considered outliers (red plusses in Fig. 4.10). These criteria can be summarized as follows:

$$\text{Outliers} \begin{cases} > Q3 + 1.5\cdot(Q3-Q1) & (4.15) \\ < Q1 - 1.5\cdot(Q3-Q1) & (4.16) \end{cases}$$

4.3.4 K-means Clustering

K-means clustering is one of several types of unsupervised classification methods. The goal of an unsupervised classification method is to find natural groupings (category labels) for a set of data without any labels, based on patterns in a similarity measure[8]. The clustering algorithm examines multiple partitions of the parameter(s) into *k* clusters in order to determine the optimal partition, such that the patterns in a cluster are more similar to each other than to patterns in different clusters. The optimum partition for a given value of k represents the solution which minimizes the sum of the within cluster variations totaled for all k clusters. The variation between each member of a specific cluster and the cluster center or centroid can be characterized by various distance measures, such as the Euclidean or squared Euclidean distance, and Mahalanobis distance.

For k-means cluster analysis, the number of clusters and the initial cluster centers need to be specified. One method to determine the initial clusters, which was used in our analysis, is to randomly divide the data into the specified number of clusters and then determine the centroid of each cluster. The final results of this

analysis are a unique cluster assignment for each sample of the data. Other variations on this technique, such as fuzzy k-means, assigns to each sample a degree of membership into the k clusters.

4.4 References

1. J. W. Tunnell, A. E. Desjardins, L. Galindo, I. Georgakoudi, S. A. McGee, J. Mirkovic, M. G. Mueller, J. Nazemi, F. T. Nguyen, A. Wax, Q. G. Zhang, R. R. Dasari and M. S. Feld, "Instrumentation for multi-modal spectroscopic diagnosis of epithelial dysplasia," *Technology in Cancer Research & Treatment*, **2**(6), 505-514 (2003).
2. S. Dreiseitl and L. Ohno-Machado, "Logistic regression and artificial neural network classification models: a methodology review," *J Biomed Inform*, **35**(5-6), 352-9 (2002).
3. D. G. Kleinbaum, M. Klein and E. R. Pryor, *Logistic regression : a self-learning text*, *Statistics for biology and health*, Springer, New York, 2002
4. A. Fielding, *Cluster and classification techniques for the biosciences*, Cambridge University Press, Cambridge, UK ; New York, 2007
5. D. G. Altman and J. M. Bland, "Diagnostic-Tests-2 - Predictive Values .4.," *British Medical Journal*, **309**(6947), 102-102 (1994).
6. T. A. Lasko, J. G. Bhagwat, K. H. Zou and L. Ohno-Machado, "The use of receiver operating characteristic curves in biomedical informatics," *Journal of Biomedical Informatics*, **38**(5), 404-415 (2005).
7. R. Mcgill, J. W. Tukey and W. A. Larsen, "Variations of Box Plots," *American Statistician*, **32**(1), 12-16 (1978).
8. A. K. Jain, R. P. W. Duin and J. C. Mao, "Statistical pattern recognition: A review," *Ieee Transactions on Pattern Analysis and Machine Intelligence*, **22**(1), 4-37 (2000).

CHAPTER 5

In Vivo Clinical Study: Healthy Volunteers

In this chapter we investigate the impact of tissue anatomy on the spectral parameters and relate the physical parameters we derive to the known tissue properties. We begin by outlining the motivation and goals for this work. Following this we describe the study protocol and methods. Finally, we end with a discussion of the results and the conclusions.

5.1 Motivation and Goals

Reflectance and fluorescence spectroscopy have been widely investigated as non-invasive tools for detecting oral cancer, and a number of studies have reported differences in spectral parameters when comparing malignant and normal tissues.[1-5] A critical step in applying these techniques to the detection of oral malignancy is to first develop an understanding of the spectral distinctions among anatomic sites (i.e. the gingiva, buccal mucosa, etc.) in the oral cavity. Many studies focus on detecting cancer in multiple sites without taking into account this natural source of variation. This can lead to spectral differences that are incorrectly attributed to malignancy.[5-8]

Only a few studies relate spectral variations to anatomic differences. In a study by Kolli *et al.*, fluorescence spectra were collected from the buccal mucosa and dorsal surface of the tongue of 21 healthy volunteers, in addition to patients.[9] For the normal sites, analysis of the ratio of fluorescence emission at 450 nm when excited at 335 and 375 nm, respectively, revealed that the intensity ratio was statistically different between the buccal mucosa and dorsal surface of the tongue. Differences in the fluorescence peak intensities at 337, 365, and 410 nm excitation were also noted by Gillenwater *et al.* in a study of 9 sites in the oral cavity of 8 healthy volunteers.[10] Another small study of 10 healthy volunteers by Dhingra *et al.* noted differences among sites.[11] Although these three studies provide support for the fact that differences based on anatomy exist, they lack a detailed description of the differences and only examine a few sites.

In a larger study by de Veld *et al.*, 9295 fluorescence spectra were collected from 97 subjects to determine whether differences in the mucosa in various sites of the mouth resulted in changes in the fluorescence intensity and lineshape.[12] A total of 7 fluorescence excitations between 350-450 nm were evaluated and emission light from 455-867 nm was analyzed. In this study both classification and clustering methods were used to identify differences by site. The authors found that the dorsal surface of the tongue and vermilion border of the lip were most consistently distinguished among the 13 sites in the mouth studied; however, significant differences in total fluorescence intensity between almost all the different sites were noted at multiple wavelengths. In this large study, the authors tested the class overlap between each of 11 anatomic sites (dorsal surface of the tongue and vermilion border of the lip were excluded because they were significantly different from rest) against the remaining 10 sites. Because the larger group consisted of multiple sites with small differences in their means and standard deviations, the resultant distribution has a large spread and is more likely to overlap with the single site against which it is compared. This would diminish the ability to observe differences. There may be added value in separating certain sites, as small changes with cancer would be more easily distinguished for narrow distributions (i.e. increased power of a diagnostic test). The authors concluded that all sites could be combined because of their considerable overlap, except the dorsal surface of the tongue and vermilion border of the lip. As shown below, we reach a different conclusion based on our own analysis.

None of the four studies cited above used a physical model of tissue fluorescence, in which the spectral features are modeled by parameters relating to

specific tissue components. Therefore, it is difficult to directly interpret or provide an explanation for the observed differences in fluorescence emission. Additionally, to the best of our knowledge, no studies examining differences in sites based on reflectance parameters have been performed.

In the present study, we use a physical model to extract parameters from both diffuse reflectance and intrinsic fluorescence spectra which are related to the morphological and biochemical properties of the tissue.[13] We compare the extracted parameter distributions for various tissue types and also relate the physical spectral parameters to the anatomic features of the tissue sites. K-means clustering is then performed in order to identify and classify groups of sites that share similar or distinct spectral properties.

Understanding the normal spectral variations in the oral cavity is essential for guiding the analysis and interpretation of data collected from lesions, and the development of robust spectroscopic-based algorithms for diagnosing oral cancer.

5.2 Materials and Methods

5.2.1 In Vivo Data Collection

Healthy volunteers (HVs) were recruited at Boston Medical Center (BMC) and at the Massachusetts Institute of Technology (MIT). The protocol for *in vivo* data collection was approved by the institutional review board at BMC and the Committee on the Use of Humans as Experimental Subjects at MIT. Written informed consent was obtained from all participants to indicate their willingness to participate in the

study. Relevant background information was obtained from each volunteer, such as their smoking history, alcohol consumption, and history of lesions in the upper aerodigestive tract. Study participants without a prior history of a lesion of the oral cavity, larynx, or esophagus (benign or malignant) were considered healthy volunteers, regardless of their smoking status. Smokers were included as HVs because in our patient study we have found that those presenting with oral lesions represent a mixture of smokers and non-smokers. Less than 1% of the subjects had a history of alcohol abuse.

Reflectance and fluorescence spectra were collected using the Fast Excitation Emission Matrix (FastEEM) instrument (Chapter 4). The optical fiber probe was disinfected with CIDEX OPA (Advanced Sterilization Products, Irvine, CA) before each session according to the manufacturer's specifications. Each measurement was acquired in approximately 1.5 s and represented the average of 5 sets of reflectance and fluorescence data. For each measurement, the standard deviation of the 5 spectra was also recorded. For each site of data collection, 3-5 measurements were acquired. We examined the raw spectra for all samples in each set of measurements for a site and excluded those with poor overlap or large standard deviations. If all the measurements in a set were inconsistent, all the data for that site was excluded.

5.2.2 Model-Based Analysis

The spectra were calibrated and the modeled fit obtained as described in Chapters 3.3 and 4.2. The diffuse reflectance spectra were fit over the range of 380-700 nm. From the reflectance spectra we derive 7 parameters: A, B, C, concentration

of hemoglobin (cHb), the oxygen saturation (α), effective vessel radius, and concentration of β-carotene ($c\beta C$).

Parameters were extracted from the fluorescence spectra in the subsequent data analysis in one of two ways, which we refer to as non-normalized fluorescence spectra or area-normalized fluorescence spectra. In the case of the non-normalized data, the basis spectrum of each fluorophore at a given excitation wavelength was normalized to the peak, such that each component had a maximum value of unity. Following this step, these basis spectra were then used to fit the intrinsic fluorescence spectrum for the sample. For area-normalization, the individual basis spectra were divided by the total area under their respective curves. This same process was applied to the intrinsic fluorescence spectrum for the sample. The resultant intrinsic fluorescence curve for the sample was then fit with a linear combination of the basis spectra. In a final step, the extracted contribution of each fluorophore derived from the fit was then divided by the total sum of the contributions for all the components at the specific excitation wavelength. The extracted contribution for each fluorophore therefore represents the fraction of the total area under the emission curve contributed by the specific fluorophore, at the specific excitation wavelength.

The fluorescence excited at 308 nm was modeled by a linear combination of tryptophan, collagen, and reduced nicotinamide adenine dinucleotide (NADH) over the range 370-600 nm. The fluorescence excited at 340 nm over the emission range 380-600 nm was modeled as a linear combination of NADH and collagen. These 2 excitations correspond to those which could be analyzed reliably from the data collected from patients (Section 4.2.1).

5.2.3 *Outlier Removal and K-means Cluster Analysis*

Outliers for each of the parameters were identified for each site separately based on the interquartile range (IQR) criterion (Section 4.3.3)[14]. Because each parameter was evaluated separately, the number of exclusions varied for each parameter. The percentage of data excluded for any one parameter was ≤13%. Following this procedure, we performed k-means cluster analysis in order to partition the entire data set into groups without assuming that each anatomic site was distinct. The analysis was performed using MATLAB (Mathworks, Natick, MA). Clustering analysis was performed after the removal of outliers in order to prevent skewing of the cluster centroid and to maximize the accuracy of the results. Analysis was performed for each parameter separately for k = 2, 3, and 4 clusters using a random initial clustering assignment and the city-block distance measure. Each separation entailed 20 replicate runs of the clustering algorithm. For each iteration, the distance from all samples to its assigned centroid was calculated and the final result of the function represented the iteration which produced the minimum sum of all the distances. For each k-means separation (i.e. k = 2, 3 or 4), the percentage of each site within each of the clusters was identified after subtracting the probability of a site being assigned to a cluster purely due to chance (i.e. 25% for k=4). We define Δ as the difference between the maximum percentage of a given site assigned to a cluster and the percentage that would be expected to be assigned due to chance. All sites for which the value for Δ for all three separations or at least two separations exceeded 30.0% were identified for each parameter. These selection criteria for Δ were used to identify sites displaying significant clustering. For each parameter, sites meeting these

criteria were either clustered together or separately based on the whether they shared the same or unique cluster assignments.

5.2.3 Impact of Probe Contact

The impact of probe pressure was tested in various sites in the oral cavity by positioning the probe on the tissue and taking three to four measurements in which the applied pressure was successively increased in between successive measurements. The probe was kept in place after each measurement. Similarly, the effect of repeat measurements (repositioning the probe in the same place) was tested by collecting three to four separate measurements in which the probe was removed and replaced in the same location in between each measurement.

5.3 Results and Discussion

5.3.1 General Description of the Data

The complete set of HV data consisted of 781 spectra. Due to instrumentation error, evidence of movement (large variations in the set if 3-5 spectra collected from each site), or too few spectra collected from a given anatomic site, 9.1% of the spectra were excluded. The final set of healthy volunteer data included 710 spectra from 79 subjects. The average age (± standard deviation) of the volunteers was 41.5±16 years. Data were collected from the following nine sites in the oral cavity: buccal mucosa (BM), dorsal surface of the tongue (DT), floor of the mouth (FM), gingiva (GI), hard palate (HP), lateral surface of the tongue (LT), retromolar trigone (RT), soft palate

(SP), and ventral surface of the tongue (VT). These sites were chosen because most of the data collected from patients were measured from these sites and because of their accessibility to the probe. Table 5.1 summarizes the number of spectra collected from each site within the oral cavity for the entire data set. Fewer spectra were collected from sites such as the HP, RT, and SP because the architecture of the oral cavity in some volunteers made these areas less accessible to the probe.

	Total Data	**Calibration Data**	**Validation Data**
Site		**Number of Spectra**	
BM	100	88	12
DT	124	110	14
FM	112	98	14
GI	58	38	20
HP	54	32	22
LT	112	101	11
RT	33	24	9
SP	43	30	13
VT	74	63	11
Total	**710**	**584**	**126**

Table 5.1 Summary of the number of spectra collected from each of the nine sites in the oral cavity for the calibration, validation, and total data sets. BM: buccal mucosa, DT: dorsal surface of the tongue, FM: floor of the mouth, GI: gingiva, HP: hard palate, LT: lateral surface of the tongue, RT: retromolar trigone, SP: soft palate, VT: ventral surface of the tongue.

A comparison of the data from each site for people who had never smoked and people who had either quit or were current smokers did not demonstrate substantial differences.

5.3.2 *Modeling the Reflectance and Fluorescence Spectra*

Modeling the Reflectance

Figure 5.1(a-b) shows representative examples of the excellent quality of the fits to the reflectance spectra. In order to fully capture the major spectral features necessary for accurately modeling the reflectance and fluorescence analysis of healthy oral tissue, we incorporated the absorber β-carotene. Although β-carotene has not been previously noted in reflectance measurements of the oral cavity, there are numerous reports in the literature of its presence in oral cells/tissue as well as in plasma.[15-17] This absorber is found at much lower concentrations than hemoglobin, but its extinction coefficient is greater than 200 times that of hemoglobin at 450 nm.[18] Distortions of the hemoglobin spectrum due to vessel packaging effects were also taken into account in the reflectance model. This correction has been used in other investigations studying tissue reflectance.[19,20]

Examples of how the quality of the modeled fit can deteriorate when either β-carotene or vessel packaging are not accounted for are shown in Figs. 5.1(a) and 5.1(b), respectively. β-carotene has a prominent absorption peak at approximately 450 nm and a secondary peak at 478 nm. The latter absorption peak is especially prominent in the spectrum in Fig. 5.1(a). The vessel packaging model predicts that the differences in the depth of the 420 nm hemoglobin peak as compared to the 540 and 580 nm peaks will diminish as the size of the effective vessel radius increases. The uncorrected hemoglobin absorption spectrum underfits the 540 and 580 nm peaks.

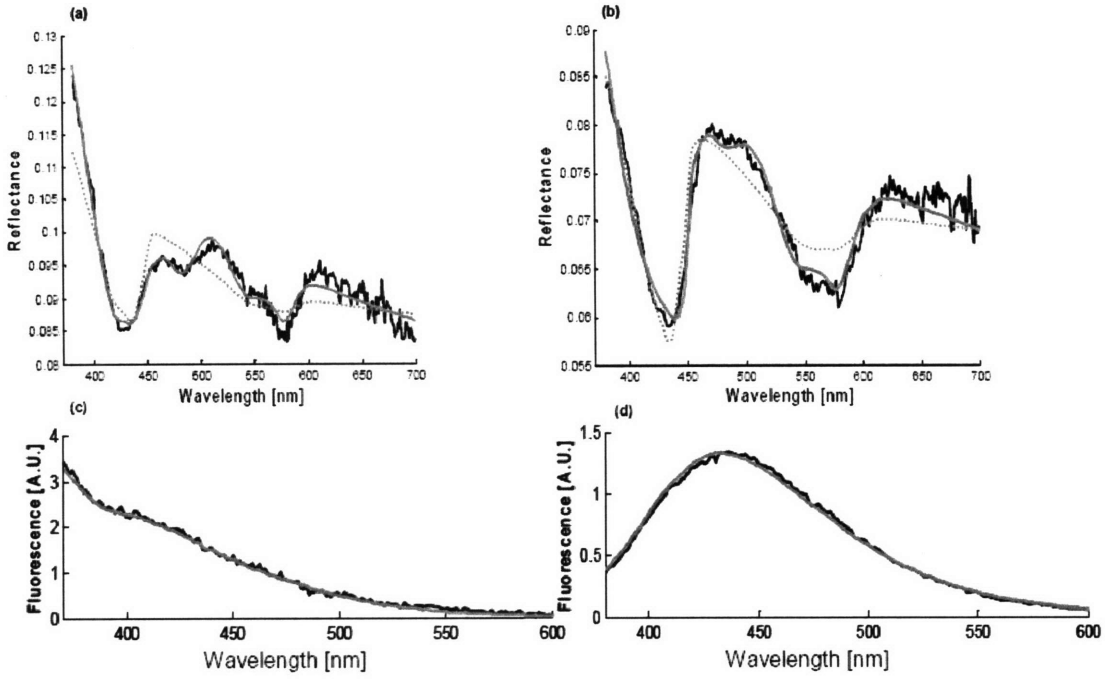

Figure 5.1 (a) Plot of a reflectance spectrum measured from the buccal mucosa (thick black line) and the modeled fit excluding β-Carotene (dashed grey line). The solid grey line demonstrates the fit when β-Carotene is included as an absorber, (b) Plot of a reflectance spectrum measured from the lateral surface of the tongue (thick black line). The grey lines show the modeled fit excluding the effect of vessel packaging (dashed grey line) and including vessel packaging (solid grey line), (c) Plot of the intrinsic fluorescence at 308 nm excitation (black line) and the modeled fit (grey line) for a DT site, (d) Plot of the intrinsic fluorescence at 340 nm excitation (black line) and the modeled fit (grey line) for a HP site.

Modeling the Fluorescence

Figure 5.1(c-d) shows examples of the fluorescence data and fit at 308 and 340 nm excitation, respectively. Figure 5.2(a) shows the spectra of tryptophan, collagen, and NADH, the three components used to fit the spectra at 308 nm excitation. A preliminary examination of the data revealed that 3 basis spectra were needed to account for the fluorescence at 340 nm excitation for the entire data set. To determine if keratin contributed to the measured fluorescence, we separated the keratinized sites (GI, HP, DT) from the non-keratinized sites and performed MCR

analysis separately on these two groups. We found that the extracted basis spectra were identical for both groups. We therefore attributed the fluorophores to collagen(s), likely in combination with elastin. The necessity for two distinct collagen spectra may result from differences in the types of collagen being excited, the layered architecture, depth-related effects (which can shift the proportions of light penetrating to the basement membrane or stromal layers), or a combination thereof. Figure 5.2(b) shows a plot of the two spectral components used to fit the spectra at 340 nm excitation in addition to NADH. We refer to the two extracted collagen components as Coll401 and Coll427, where the number represents the wavelength at which the peak emission occurs. Figure 5.3 shows an example of the deviation in the fit to the fluorescence data at 340 nm excitation that was systematically observed throughout the data when only 2 basis spectra were used to fit the data [Figure 5.3(a)]. Figure 5.3(b) shows the fit quality when NADH, Coll401, and Coll427 were used to fit the data.

To determine if the two fluorophores we extracted could be due to different types of collagen we collected the fluorescence emission with the FastEEM from dry collagen IV (Sigma-Aldrich C9879) and collagen I (Sigma-Aldrich C5533), the major forms present in the basement membrane and stroma, respectively.[21,22] Figure 5.2(c) shows the fluorescence emission excited at 340 nm. The peaks for collagen IV and collagen I occur at ~422 and ~409 nm, respectively.

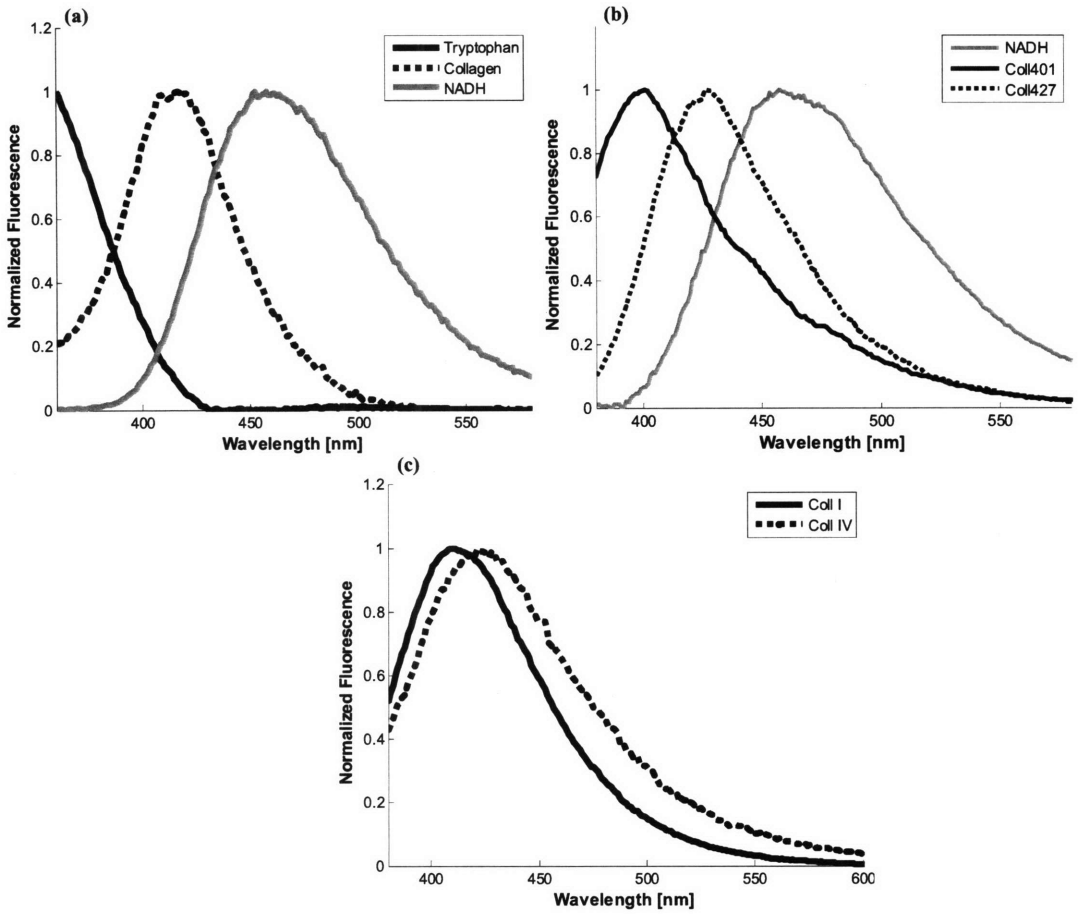

Figure 5.2 Basis spectra used to fit the fluorescence spectra (a) Plot of the basis spectra used at 308 nm excitation: tryptophan, Collagen, and NADH (b) Plot of the two fluorescence spectra extracted at 340 nm excitation using multivariate curve resolution, Coll401 (solid black line) and Coll427 (dotted black line), and NADH (shaded grey line). The fluorescence peaks occur at approximately 401 nm and 427 nm, respectively. These three components were used to model the intrinsic fluorescence spectra at 340 nm excitation. (c) Plot of the fluorescence emission from collagen I (solid line) and collagen IV (dotted line) chemicals at 340 nm excitation. The fluorescence peaks occur at approximately 409 nm and 422 nm, respectively.

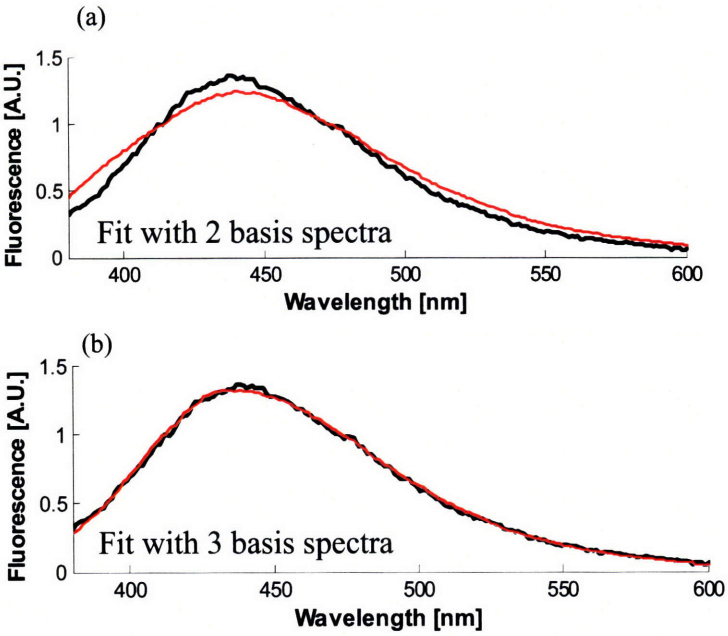

Figure 5.3 (a) Example of the deviation in the fit at 340 nm excitation observed when only 2 basis spectra were used to fit the data. (b) Improvement in the quality of the fit when three basis spectra are used (NADH, Coll401, and Coll427). In each plot the data is shown in black and the fit in red.

The spectra collected from the two collagens show that collagen IV (major type in the basement membrane) has a different emission profile from collagen I (major type in the stroma)[23]. This supports the fact that variations in the collagen signal exist, although the emission of the extracted spectral components and chemicals do not appear to be equivalent. A study of freshly excised *ex vivo* cervical tissue by Chang *et al.* found that the stromal fluorescence differed greatly among various specimens and that data from their large patient data set was best modeled using a linear combination of three separate collagen fluorescence spectra[24]. The authors attributed two of the components to different types of collagen crosslinks, as they are the structural components responsible for collagen fluorescence. Differences between the spectra of the pure chemicals and the results from MCR are likely due to the inability of the dry chemical to recapitulate the actual protein *in situ* and the influence of the tissue microenvironment, the contribution of other types of collagen,

and fluorophores such as elastin and possibly keratin, which strongly overlap in emission.

The following 6 fluorescence ratios were examined based on the analysis of the non-normalized data: NADH/Collagen (308 nm), NADH/Tryptophan (308 nm), NADH/Coll401 (340 nm), NADH/Coll427 (340 nm), NADH/Total collagen (340 nm), and Coll401/Coll427 (340 nm). The number in parentheses indicates the excitation wavelength. Total collagen refers to the sum of Coll401 and Coll427. Analysis of the area-normalized intrinsic fluorescence yielded 6 additional parameters: Tryptophan (308 nm), Collagen (308 nm), NADH (308 nm), NADH (340 nm), Coll401 (340 nm), and Coll427 (340 nm).

5.3.3 Model-based Analysis Parameter Distributions

Table 5.2 summarizes the mean, standard deviation (σ), and the relative standard deviation (RSD) ([σ/mean]*100) for all the parameters examined for each of the 9 individual sites, as well as when all sites are combined into a single group. In each case the statistical measures were calculated after applying the IQR exclusion criteria. No clear trend was noted among the excluded samples, therefore effects such as probe contact (i.e. slipping of the probe against the tissue, probe pressure, etc.) and tissue movement likely are the cause of these outliers.

Parameter	BM	DT	FM	GI	HP	LT	RT	SP	VT	Combined
A										
n	95	120	109	55	51	108	33	42	72	685
Mean	1.7	1.4	1.5	1.5	1.4	1.5	1.7	1.6	1.5	1.5
σ	0.1	0.1	0.2	0.2	0.2	0.1	0.2	0.1	0.1	0.2
σ/Mean [%]	8.6	8.8	14	12	11	7.9	11	9.4	9.1	12
B										
n	100	124	110	51	52	110	33	43	74	697
Mean	0.4	0.3	0.2	0.6	0.6	0.1	0.4	0.3	0.2	0.3
σ	0.2	0.2	0.1	0.1	0.2	0.1	0.2	0.2	0.1	0.2
σ/Mean [%]	47	48	65	20	28	69	48	54	68	64
C										
n	98	121	110	55	51	109	33	42	74	693
Mean	0.010	0.012	0.02	0.004	0.002	0.018	0.008	0.009	0.02	0.013
σ	0.006	0.007	0.01	0.004	0.002	0.006	0.008	0.005	0.01	0.009
σ/Mean [%]	61	62	53	96	112	32	96	59	46	72
cHb										
n	88	113	103	53	48	102	32	38	72	649
Mean	1.2	1	2	0.6	0.4	3	1.0	0.6	2	1
σ	0.6	1	1	0.3	0.1	1	0.8	0.3	1	1
σ/Mean [%]	45	70	80	58	29	55	81	48	61	82
α										
n	99	123	111	56	53	112	33	42	74	703
Mean	0.6	0.7	0.6	0.6	0.7	0.7	0.6	0.7	0.6	0.6
σ	0.1	0.1	0.2	0.1	0.1	0.1	0.2	0.1	0.2	0.2
σ/Mean [%]	24	20	32	16	13	22	25	21	29	24
Vessel Radius [mm]										
n	96	115	106	51	49	106	32	39	74	668
Mean	0.004	0.005	0.005	0.0032	0.0028	0.006	0.004	0.0030	0.006	0.005
σ	0.001	0.002	0.002	0.0006	0.0002	0.002	0.002	0.0003	0.002	0.002
σ/Mean [%]	31	40	46	18	7	33	44	11	36	44
cβC										
n	96	123	107	55	53	106	31	43	73	687
Mean	0.003	0.002	0.003	0.0012	0.0014	0.003	0.002	0.004	0.003	0.003
σ	0.001	0.001	0.002	0.0008	0.0008	0.001	0.001	0.002	0.002	0.002
σ/Mean [%]	45	71	57	64	58	46	50	48	54	65
NADH/Collagen (308 nm)										
n	100	120	102	56	46	105	33	41	73	676
Mean	0.4	0.39	0.6	0.34	0.40	0.39	0.3	0.5	0.5	0.4
σ	0.1	0.09	0.1	0.08	0.06	0.09	0.1	0.1	0.1	0.1
σ/Mean [%]	30	23	17	24	15	23	36	24	21	29
NADH/Tryptophan (308 nm)										
n	98	119	110	51	52	107	33	39	72	681
Mean	0.22	0.14	0.4	0.08	0.12	0.18	0.13	0.25	0.3	0.2
σ	0.08	0.05	0.1	0.03	0.04	0.06	0.07	0.08	0.1	0.1
σ/Mean [%]	36	39	30	38	31	32	59	30	38	58

Parameter	BM	DT	FM	GI	HP	LT	RT	SP	VT	Combined
NADH/Coll401 (340 nm)										
n	100	118	109	54	51	109	33	41	73	688
Mean	0.6	0.8	0.8	1.1	1.3	0.4	0.5	0.8	0.7	0.8
σ	0.2	0.3	0.2	0.6	0.4	0.2	0.2	0.2	0.2	0.4
σ/Mean [%]	32	36	27	53	29	44	46	29	27	49
NADH/Coll427 (340 nm)										
n	98	120	106	52	49	108	31	42	71	677
Mean	0.6	0.6	0.6	0.7	0.8	0.5	0.6	0.6	0.5	0.6
σ	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.1	0.2
σ/Mean [%]	43	37	38	22	22	42	30	27	25	38
NADH/Total Collagen (340 nm)										
n	99	117	109	58	49	109	32	42	70	685
Mean	0.3	0.4	0.3	0.4	0.5	0.22	0.2	0.4	0.28	0.3
σ	0.1	0.1	0.1	0.2	0.1	0.09	0.1	0.1	0.06	0.1
σ/Mean [%]	36	31	31	36	23	42	39	28	22	40
Coll401/Coll427 (340 nm)										
n	96	120	106	57	50	111	30	42	71	683
Mean	0.9	0.8	0.8	0.7	0.6	1.1	1.2	0.8	0.8	0.8
σ	0.2	0.3	0.2	0.3	0.1	0.3	0.3	0.1	0.1	0.3
σ/Mean [%]	27	34	25	41	19	29	25	16	19	34
Tryptophan (308 nm)										
n	100	120	109	51	54	108	33	39	72	686
Mean	0.28	0.35	0.20	0.46	0.39	0.30	0.37	0.25	0.23	0.3
σ	0.06	0.07	0.04	0.06	0.06	0.05	0.09	0.05	0.06	0.1
σ/Mean [%]	23	20	19	12	14	17	24	19	25	32
Collagen (308 nm)										
n	91	114	103	55	49	104	32	39	70	657
Mean	0.43	0.40	0.43	0.37	0.37	0.44	0.42	0.42	0.44	0.42
σ	0.04	0.03	0.03	0.03	0.02	0.03	0.03	0.03	0.03	0.04
σ/Mean [%]	8.8	8.2	7.9	7.2	5.3	7.6	7.1	8.1	5.7	9.5
NADH (308 nm)										
n	100	121	106	56	52	109	33	38	74	689
Mean	0.28	0.24	0.36	0.19	0.23	0.26	0.21	0.31	0.32	0.27
σ	0.06	0.05	0.05	0.06	0.04	0.05	0.07	0.05	0.06	0.08
σ/Mean [%]	22	22	13	31	19	20	36	15	20	28
NADH (340 nm)										
n	100	118	106	58	49	109	32	42	73	687
Mean	0.29	0.34	0.32	0.37	0.40	0.23	0.27	0.32	0.28	0.31
σ	0.08	0.07	0.06	0.08	0.05	0.08	0.07	0.06	0.05	0.08
σ/Mean [%]	27	21	19	23	13	33	27	19	18	27
Coll401 (340 nm)										
n	100	122	111	58	52	110	32	39	74	698
Mean	0.32	0.28	0.29	0.25	0.22	0.38	0.39	0.30	0.30	0.30
σ	0.05	0.06	0.04	0.09	0.04	0.07	0.07	0.03	0.04	0.07
σ/Mean [%]	16	23	14	35	18	19	19	11	15	25
Coll427 (340 nm)										
n	92	118	103	54	54	104	32	43	71	671
Mean	0.37	0.38	0.39	0.38	0.37	0.37	0.34	0.37	0.41	0.38
σ	0.06	0.07	0.06	0.04	0.05	0.05	0.05	0.04	0.04	0.06
σ/Mean [%]	16	18	16	11	13	13	14	10	10	15

Table 5.2 (Continued from previous page) Summary of the mean, standard deviation, and relative standard deviation for each parameter for each of the nine sites, as well as the combined group of sites. BM: buccal mucosa, DT: dorsal surface of the tongue, FM: floor of the mouth, GI: gingiva, HP: hard palate, LT: lateral surface of the tongue, RT: retromolar trigone, SP: soft palate, VT: ventral surface of the tongue.

Scattering Parameters

For the A parameter, most of the sites showed very similar mean values, with the exceptions of the BM and RT. The BM and RT exhibited a mean which was significantly higher than all other sites based on a test of multiple corrections ($p<0.05$), but not from each other. The HP and DT demonstrated slightly lower mean values which were significantly different from all other sites except the GI.

The GI and HP both exhibited values of B of 0.6, which is significantly larger than for other sites. Three groups were easily distinguished for the C parameter. The FM, LT, and VT formed one group, demonstrating the largest values of C. The BM, DT, RT, and SP form another group with intermediate values of the C parameter. At the other extreme, the keratinized GI and HP are easily distinguished by a small C parameter. Figure 5.4(a) shows an example of a box plot showing the C parameter distribution for each of the 9 sites.

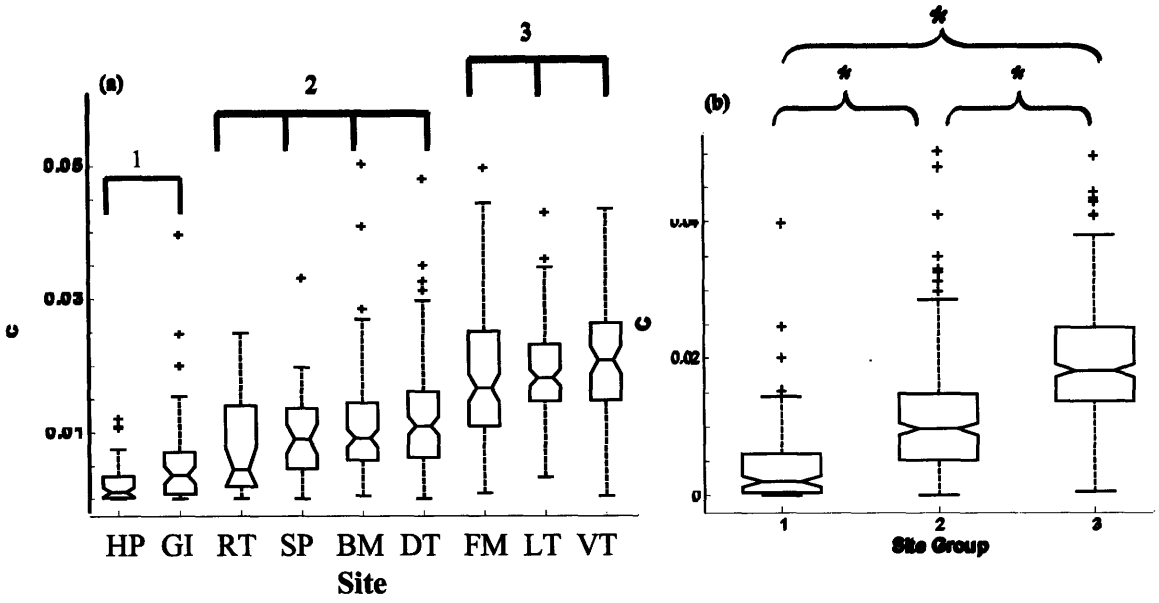

Figure 5.4 (a) Example of the distribution for the C parameter obtained for each of the 9 sites. The brackets indicate the 3 groups of sites which are similar in their mean values. (b) A box plot which shows the result of combining the individual sites shown by the brackets into three site groups. The asterisks indicate a significant difference between two groups.

It is clearly seen that both the median and spreads in the distribution vary among the 9 sites, however, there are approximately 3 distinct groups of sites (site groups) and the members of each group share similar properties. When the members of each of the three site groups are combined in their respective group, the resultant distributions are shown in Figure 5.4(b). The means between all three site groups are statistically different from one other.

Absorption Parameters

The GI, HP, and SP demonstrated very low concentrations of hemoglobin as compared to other sites. In the case of the former two sites, these results may indicate that keratin prevents light from reaching the underlying stroma where the blood

vessels are located. During data collection from the GI, visible blanching was often observed when the probe contacted the tissue, which may also contribute to the lower cHb values obtained for this site. The FM, LT, and VT, sites with relatively thin epithelia as compared to other sites, demonstrated the highest cHb. The mean oxygen saturations were comparable across all 9 sites and the combined set of sites had a mean value of 60%. The mean effective vessel radius for the nine sites ranged from approximately 3-6 μm. The mean effective vessel radius was largest for areas on the tongue (DT, VT, LT) and the FM. Blood vessels are easily observable at the macroscopic level on the VT and FM and the tongue has a very high density of vessels.

Because β-carotene is present in adipose tissue and in blood, we would expect the sites which demonstrated a large cHb and contain submucosal fat to contain the largest concentration of this absorber. The highest concentration was found in the SP, which meets both these criteria. The FM, LT, and VT and BM were tied for the next highest value. The former three sites were those which displayed the highest cHb and the BM and FM both contain adipose tissue in the submucosa. Similar to the results for cHb, the GI and HP exhibit the lowest concentration of the absorber.

Fluorescence Parameters

For the NADH/Tryptophan and NADH/Collagen fluorescence ratios at 308 nm excitation, the FM was significantly different from all other sites. At 308 nm excitation, light does not penetrate deeply into tissue because of absorption, and the emission signal will be heavily influenced by superficial components. The FM

displayed a mean value significantly higher ($p<0.05$) than all the other sites for the NADH/Tryptophan (308 nm) and NADH/Collagen (308 nm) ratios, as well as the NADH (308 nm) and Tryptophan (308 nm) parameters. The FM is distinguished by a very thin epithelium (190 ± 40 μm), as compared to 580 ± 90 μm for the BM, and 310 ± 50 μm for the HP[21]. The GI and HP displayed the two lowest mean Collagen (308 nm) values.

Of the two collagen fluorophores excited at 340 nm excitation, the contribution of fluorescence emission peaking at 427 nm was slightly higher than the contribution from the peak at 401 nm for all of the sites except the LT and RT. The site with the largest values for the ratio of NADH to Coll401, Coll427, and the Total collagen contribution was the HP. The GI had the second highest value for the former two ratios, but was tied for second with the DT and SP for the latter ratio. These ratios are affected by a complex mixture of epithelial and stromal characteristics. The GI and HP demonstrated the lowest mean Coll401/Coll427 ratio of all the sites. In a study by Collier *et al.* examining cervical tissue by confocal microscopy, the authors found that keratin (which can be present both in normal and precancerous tissue) introduces a significant source of variability in the depth penetration of the incident light, and therefore the extent to which scattering from different layers contribute to the observed signal.[25] Although the collagen content in the deeper stroma (mostly collagen type I) is significantly higher, this is offset by the higher probability of light reaching and returning from more shallow structures, such as the collagen IV present in the basement membrane. Most sites display Coll401/Coll427 ratios less than one, indicating that emission from Coll427 is preferentially collected over Coll401. The

presence of keratin on the GI and HP likely further reduces penetration of light to the deeper stroma thus resulting in the low Coll401 (340 nm) values, and as a result, mean Coll401/Coll427 values.

General Trends

The detailed summary of the parameters provided in Table 5.2 reveals a number of findings relating the extracted parameters to the structural characteristics of the sites. The GI, DT, and HP are sites which are normally keratinized, however only the GI and HP display unique features as compared to the non-keratinized sites. The keratinized sites are distinguished by large values for B, Tryptophan (308 nm), and NADH (340 nm), and low C, cHb, effective vessel radius, cβC, Coll401/Coll427, NADH/Total collagen, Collagen (308 nm), and NADH (308nm) values. The large B value for the GI and HP may reflect strong scattering by the superficial keratin layer, resulting in a rapid fall in the scattering as a function of wavelength. In their depth-resolved studies of fluorescence from epithelial tissue, Wu *et al.* observed a similar trend in the fluorescence intensity.[26] Unlike for non-keratinized tissues, they found that the presence of keratin caused the fluorescence intensity to rapidly decay with increasing sampling depth. The DT, which is only partially keratinized, exhibited a B value comparable to the non-keratinized sites.[27] The presence of keratin appears to diminish signals from small, more superficial cellular scatterers, resulting in small C values for the HP and GI.

The relative standard deviation varies widely across sites for a given parameter even among sites with similar means. For 16 of the 19 parameters, the RSD

for the combined group of sites exhibits a value which ranks among the highest 3 values for a given parameter. Figure 5.5 shows examples of the distributions for two parameters, the effective vessel radius and NADH/Coll427 (340 nm). The HP was frequently shown in Table 5.2 to have a relatively small RSD compared to the other sites, and the combined group to have among the largest values.

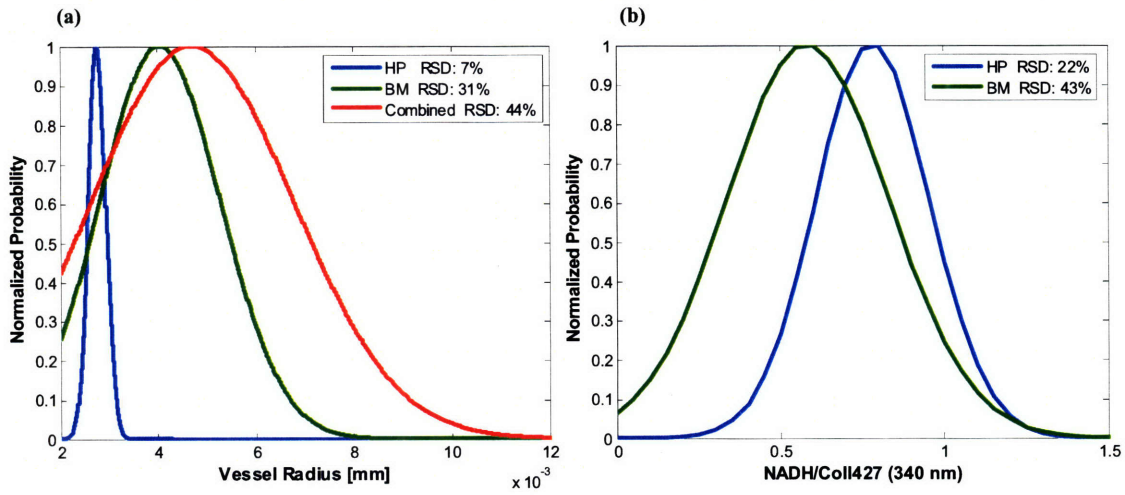

Figure 5.5 (a) Example of the distributions for the effective vessel radius for the HP (blue line), BM (green line), and the combination of 9 sites (red line). (b) Example of the distributions for the NADH/Coll427 ratio for the HP (blue line) and the combination of 9 sites (green line).

In fact for most parameters, HP and, to a lesser extent the SP, were also unique in that the spread in the distribution for the samples was significantly smaller than for most other sites. This may be due to the relatively small number of samples/subjects for these sites as compared to other sites as there may be inter-subject variability. Another possible reason is that these sites are relatively hard and non-compressible and therefore less prone to error from variations in probe pressure.

The fluorescence parameters extracted from the area-normalized fluorescence may be more useful in future studies focused on disease detection, as they exhibited

smaller spreads in their distributions than for the parameters extracted from the non-normalized fluorescence spectra.

In addition to the natural variation across various sites as a result of the architecture of the mucosa, regional variations in the morphology within a given site, poor accessibility with the probe (i.e. the RT, some locations along the palate), the intrinsic movement of the tissue (i.e. in the case of the tongue), and the highly compressible nature of the softer tissue sites also contribute to the spread in the data. Analysis of data collected at different points along the palate and tongue demonstrated not only the variation within a site, but also that transition areas between sites can produce either gradual or marked shifts in a parameter, depending on its sensitivity to the specific structural differences present between the two sites. These transitions can be appreciated based on the clinical appearance and through microscopic examination of the oral cavity.[21]

The sources of the variations in the data will be discussed in further detail in the subsequent discussion of results for the validation data test (Section 5.3.5 and 5.3.6).

5.3.4 K-means Cluster Analysis

K-means clustering (Section 4.3.4) was used to determine if there were specific sites with unique properties without assuming *a priori* that the 9 individual sites were well-defined groups (unsupervised classification). This analysis provided an objective and quantitative means of identifying and determining the magnitude of similarities and differences among sites based on the data summarized in Table 5.2.

In applying this method, all of the samples used to generate the data summarized in Table 5.2 were inputted into the algorithm. The algorithm assigned each of the samples to a specific cluster based on the parameter value, such that samples with similar parameter values are grouped together. The algorithm does not incorporate any information about the specific anatomic site from which the sample was derived. We then determined what percentage of the samples from each anatomic site were assigned to each cluster and applied the specific criteria for Δ (Section 5.2.3) to identify similar and dissimilar groups of sites. The various site clusters are shown in Table 5.3. For each parameter, we list the site(s) meeting the criteria for Δ and the range of values for Δ for each cluster. For those sites meeting these criteria, each row represents a separate clustering assignment; therefore, sites which were assigned to the same cluster are listed together, while those assigned to a separate cluster are listed in different rows. The sites which are not listed for a given parameter are not necessarily all assigned to the same cluster.

For all but one parameter, oxygen saturation, there is at least one site or group of sites that form a well-defined cluster. For 15 of the 19 parameters studied, the HP was identified as being distinct from the majority of the other sites, either alone or in combination with another site(s). The next site to most frequently demonstrate significant clustering was the GI (10 parameters). For 8 of the 19 parameters, the HP was assigned to the same cluster as the GI. Only in the case of one parameter (NADH/Tryptophan (308 nm)), did the DT group with these sites. The FM displayed unique properties for all of the non-normalized fluorescence parameters at 308 nm excitation, and 2 of the 3 area-normalized fluorescence parameters. The sites least

likely to show significant clustering were the BM and DT. Each of these sites met the criteria for Δ only for a single parameter.

Sites meeting this criteria exhibit a very narrow distribution about their mean and/or a very distinct mean value as compared to the rest of the sites. In support of this, the HP, as shown in Table 5.2 exhibits small RSDs and mean parameter values which are frequently at the extreme of the range for the 9 sites. Although the GI and HP frequently cluster together, the DT only displays significant clustering with these sites for a single parameter (NADH/Tryptophan (308 nm)). In addition to differences in the extent of keratinization, the DT is distinct from the GI and HP because it contains muscle in the submucosa rather than bone. It can be noted from the range of values for Δ observed for a number of parameters, that the degree of clustering can be very significant.

Parameter	Sites	Range for Δ [%]
A	BM	30.9-38.4
B	GI,HP	36.5-56.1
	LT,VT	30.4-44.5
C	GI,HP	38.6-66.7
cHb	GI,HP,RT,SP	34.4-75.0
	BM	30.7-40.9
α	No distinct sites	
Vessel Radius	BM,GI,HP,SP	32.3-75.0
cβC	GI,HP	38.4-48.2
NADH/Collagen (308nm)	GI	30.4-41.1
	HP	31.5-34.1
	FM	32.4-48.0
NADH/Tryptophan (308nm)	DT,GI,HP,RT	34.9-69.1
	FM	36.7-43.6
NADH/Coll401 (340nm)	HP	42.2-47.5
	LT,RT	34.8-42.7
NADH/Coll427 (340nm)	LT,VT	30.6-39.8
NADH/Total Collagen (340nm)	HP	39.8-50.5
	LT	33.6-38.1
Coll401/Coll427 (340nm)	HP	48.0-53.7
	RT	40.0-41.7
Tryptophan (308nm)	GI,HP	35.2-65.2
	FM,SP,VT	30.8-63.9
Collagen (308nm)	GI,HP	46.4-56.5
NADH (308nm)	HP	34.6-34.6
	GI	37.5-47.0
	FM	40.1-53.5
	SP	34.2-40.8
NADH (340nm)	HP	39.8-50.3
	LT	30.9-36.5
Coll401 (340 nm)	HP	48.1-57.1
	RT	34.4-40.6
Coll427 (340nm)	VT	34.3-35.9

Table 5.3 Summary of the results from the k-means clustering analysis. For each parameter, cases in which Δ exceeded 30.0% for all three separations (k=2, k=3, and k=4) or at least two separations are listed. For sites meeting these criteria, those assigned to the same cluster are listed together, while sites assigned to different clusters are listed separately. In the last column, the range of values for Δ observed for a specific cluster is provided. BM: buccal mucosa, DT: dorsal surface of the tongue, FM: floor of the mouth, GI: gingiva, HP: hard palate, LT: lateral surface of the tongue, RT: retromolar trigone, SP: soft palate, VT: ventral surface of the tongue.

The results from the K-means cluster analysis provided in Table 5.3 provide a rapid means of determining which parameters would be expected to play a significant role in distinguishing a certain site or group of sites from one another. In Figure 5.4 we showed that there were several groups of sites which displayed similar properties for the C parameter. In Table 5.3 the GI and HP were found to cluster significantly (met the criteria for Δ) for this parameter and are listed as belonging to the same cluster. The sites in the remaining two groups have large spreads, and therefore were not found to cluster significantly according to the established criteria.

Based on the information provided in Table 5.3, the BM and FM would be expected to show some degree of separation based any of the following parameters: A, cHb, vessel radius, NADH/Collagen (308 nm), NADH/Tryptophan (308 nm), Tryptophan (308 nm), and NADH (308 nm). In each case either the BM and/or FM met the criteria for Δ, and if both meet the criteria, they were assigned to separate clusters. Figure 5.6 shows box plots for two of the parameters, A and NADH/Tryptophan (308 nm). A shift in the medians for the 2 sites is apparent for both parameters.

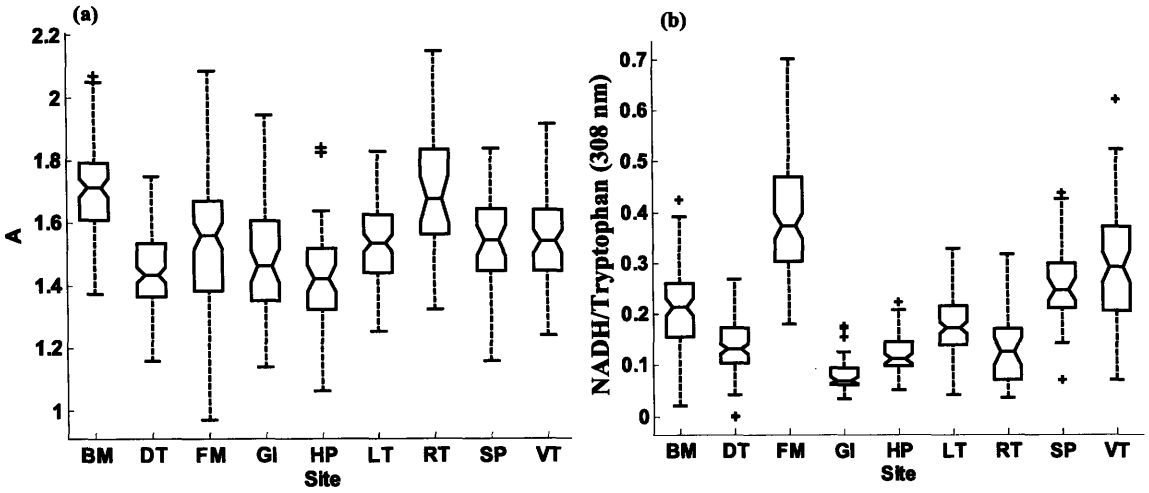

Table 5.6 Boxplots showing the distributions for each of the 9 sites for (a) the A parameter, and (b) NADH/Tryptophan (308 nm).

Figure 5.7(a) shows a binary scatter plot demonstrating a clear separation for these sites based on these two parameters. Decision lines were generated using logistic regression and the sensitivity and specificity were evaluated using leave-one-out cross validation. The sensitivity and specificity for separating the BM from the FM using only the 2 parameters (NADH/Tryptophan (308 nm) and A) were 81% and 83%, respectively (ROC-AUC=0.89). Similar results could also be obtained using either Tryptophan (308 nm) and A (sensitivity and specificity of 87% and 79%, respectively) or NADH (308 nm) and A (sensitivity and specificity of 83% and 78%, respectively).

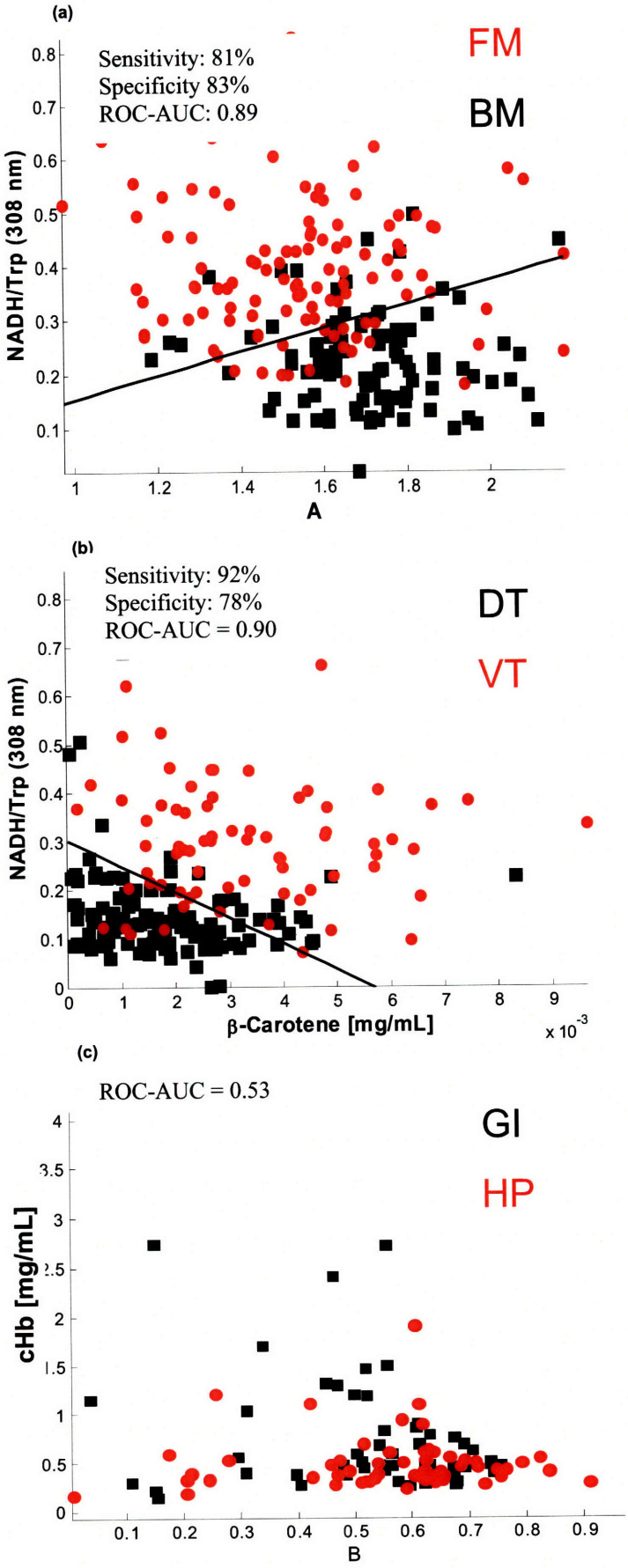

Figure 5.7 Binary scatter plots comparing two sites. (a) A plot of NADH/Tryptophan (308 nm) versus A for the FM (grey squares) and BM (black circles), (b) A plot of NADH/Tryptophan (308 nm) versus cβC for the DT (grey circles) and VT (black squares), and (c) A plot of cHb versus B parameter for the GI (black squares) and HP (grey circles). The decision lines derived from selecting the greatest combined sensitivity and specificity for separating the two sites are shown in (a) and (b).

Another binary plot comparing the DT and VT is shown in Figure 5.7(b) for NADH/Tryptophan (308 nm) versus cβC. The two different aspects of the tongue can be differentiated based on NADH/Tryptophan (308 nm) and cβC with a sensitivity and specificity of 92% and 78%, respectively (ROC-AUC=0.90). As expected from Table 5.3, NADH/Tryptophan (308 nm) contributed to the separation to a greater extent than the cβC. Using the parameters extracted after area-normalization, similar results were demonstrated. A sensitivity and specificity of 92% and 81%, respectively, was obtained using cβC and Tryptophan (308 nm). A sensitivity and specificity of 89% and 79%, respectively, was obtained using cβC and NADH (308 nm).

Figure 5.7(c) shows a similar binary scatter plot of cHb versus B for the GI and HP. Both of these sites were found to cluster significantly and were assigned to the same cluster for these two parameters. As the plot shows, the two sites overlap considerably (ROC-AUC=0.53).

Despite the large RSD values and apparent overlap among the sites, clear separations based on anatomy can occur.

Figure 5.8 shows the similarity of the GI and HP histology [Figs. 5.8(a) and 5.8(b)], and their contrast to the thicker, non-keratinized epithelium of the BM [Fig. 5.8(c)]. A multiple comparison test revealed that the BM and GI were significantly different for 15/19 parameters, and the BM and HP were significantly different for 16/19 parameters. With only two parameters (cHb and A), the BM and HP could be distinguished with a sensitivity and specificity of 89% and 95%, respectively (ROC-AUC = 0.93).

Although the separation between specific sites may be small for a single parameter, the combination of several parameters can enhance their individual effects and produce a considerable separation. By combining C with A and Tryptophan (308 nm), a sensitivity and specificity of 94% and 80%, respectively, can be achieved in discriminating between the BM and FM (ROC-AUC = 0.88 versus ROC-AUC = 0.84). As another example, the FM and VT are poorly distinguished (sensitivity and specificity of 80% and 47%, respectively) based on the A parameter and Tryptophan (308 nm) (ROC-AUC = 0.68). However, with the addition of NADH (340 nm), a sensitivity and specificity of 72% and 72%, respectively, can be achieved (ROC-AUC = 0.73).

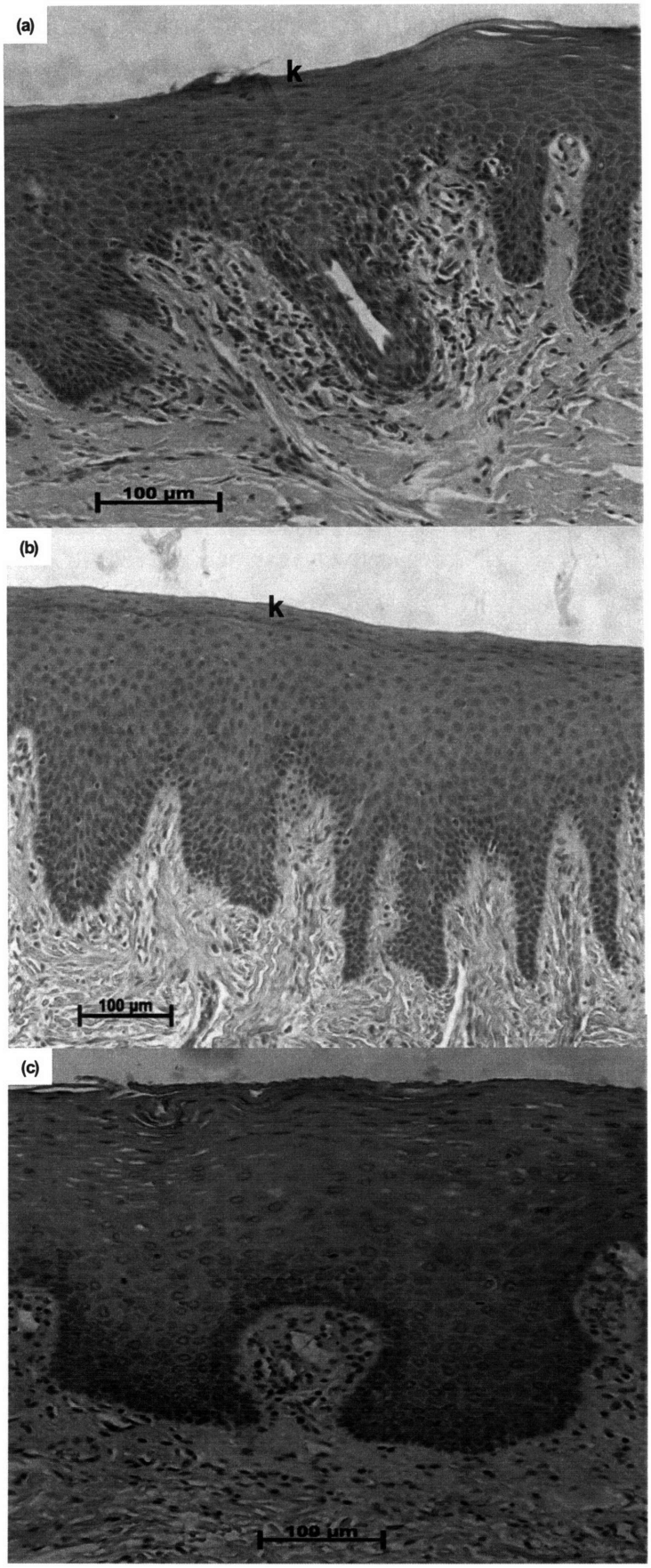

Figure 5.8 Photos showing the normal microscopic architecture of the HP (a) and the GI (b), sites which were frequently found to be spectroscopically similar to each other. The contrasting architecture of the BM is shown in (c). Overlying keratin is indicated by the letter, k.

5.3.5 *Validation of Normal Tissue Discrimination*

In order to evaluate how accurately we have captured the parameter distributions for normal tissue ("normal limits") for each site and can use the results in Table 5.2 to characterize unknown data, we divided the full data set into two groups: a calibration data set, and a validation data set. The calibration data set consisted of 584 spectra collected from 65 subjects and was used to derive the normal limits for each parameter for each site based on the 95% confidence interval (C.I.). The IQR exclusion criterion was applied to the calibration data in a similar manner as for the previous analysis of the full data set. The validation data was comprised of 126 spectra collected from the remaining 14 subjects. This data set was used to test the reliability of the calibration data to correctly categorize normal tissue by determining if each parameter extracted for a given site fell within the 95% C.I. for that parameter for that specific site. The average age (± standard deviation) for subjects in the calibration and validation data was 41.0±16 and 43.8 ±18 years, respectively. Table 5.1 provides a summary of the number of spectra collected from each site in both data sets. The percentage of all the samples in the validation data falling within the normal limits developed from the calibration data set (after applying the IQR outlier criterion) is shown in Table 5.4. The number of correctly classified samples in the validation data for each site and for each parameter was determined. We then totaled the sums across all the sites for each parameter and divided by the total number of samples to determine the percentage of correctly classified samples.

Extracted Parameter	Percentage of Data within 95% C.I.
A	95.2
B	93.7
C	93.7
cHb	88.1
Oxygen Saturation	95.2
Vessel radius	92.1
cβC	88.1
NADH/Collagen (308 nm)	96.8
NADH/Tryptophan (308 nm)	93.7
NADH/Coll401 (340 nm)	92.1
NADH/Coll427 (340 nm)	88.9
NADH/Total Collagen (340 nm)	92.9
Coll401/Coll427	97.6
Tryptophan (308 nm)	93.7
Collagen (308 nm)	93.7
NADH (308 nm)	96.0
NADH (340 nm)	94.4
Coll401 (340 nm)	95.2
Coll427 (340 nm)	96.8

Table 5.4 Summary of the percentage of the 126 spectra in the validation data collected from healthy volunteers in which the listed parameters fell within the 95% confidence interval ('normal limits') for a specific parameter for a given site. The normal limits were developed based on the results of the calibration data set.

We obtained excellent results, with few false positives (abnormal results), indicating that we have successfully captured the distributions for each parameter. All of the fluorescence parameters except (NADH/Coll427) demonstrate excellent accuracy in classifying normal tissue. The concentrations of hemoglobin and β-carotene are somewhat less accurate for identifying normal tissue. This is not surprising in that within any given site, (particularly those in which vessels may be

located non-uniformly throughout the site), the probe may sample more or less hemoglobin or β-carotene depending on the proximity and density of the blood vessels.

5.3.6 Impact of Probe Contact

There are a number of factors that contribute to the spread observed for a given site for the extracted parameters. These include variations among and within individuals, variations in the anatomy within a site, as well as probe-related factors. Since the healthy volunteer data were collected from a large number of subjects, the effect of probe contact likely contributes to the observed spread for each site. In order to examine the impact of the probe we collected sets of 3-4 measurements in 13 people in which we either repeated a measurement in the same location, or increased the pressure on the probe in between measurements. Nine sets of data from the BM, DT, FM, LT, and VT were collected in which the probe pressure was varied (29 spectra total). Twelve sets of data were collected from the BM, DT, FM, GI, HP, and LT in which the probe was repositioned in the same site in between successive measurements (36 spectra total). We applied the 95% CI criteria (normal limits) developed from the data summarized in Table 5.2 to test whether the parameters derived from the data in each of these 2 tests were classified as normal tissue. The comparison against the normal limits was performed in a site specific manner. The results are shown in Table 5.5.

Parameter	Repeat Measurements [%]	Pressure Varied [%]
A	91.7	96.6
B	94.4	100.0
C	91.7	96.6
cHb [mg/mL]	72.2	100.0
α	94.4	86.2
Vessel radius [mm]	88.9	100.0
cβC	80.6	93.1
NADH/Collagen (308 nm)	80.6	100
NADH/Tryptophan (308 nm)	97.2	100
NADH/Coll401 (340 nm)	94.4	82.8
NADH/Coll427 (340 nm)	100	100
NADH/Total Collagen (340 nm)	100	82.8
Coll401/ Coll427 (340 nm)	100	86.2
Tryptophan (308 nm)	91.7	96.6
Collagen (308 nm)	83.3	89.7
NADH (308 nm)	91.7	100.0
NADH (340 nm)	100.0	86.2
Coll401 (340 nm)	100.0	86.2
Coll427 (340 nm)	100.0	100.0

Table 5.5 Percentage of measurements falling within the 95% confidence intervals developed using data collected from normal tissue for repeat measurements in the same site and when the pressure was varied.

The impact of probe pressure is particularly pronounced for the fluorescence parameters, particularly at 340 nm excitation, while repeating a probe measurement in the same location produces misclassifications mostly for reflectance parameters and for fluorescence parameters at 308 nm excitation. For 8/9 sets, NADH/Coll427 (340 nm), NADH/Total collagen (340 nm), Tryptophan (308 nm), and NADH (340 nm) consistently decreased, and NADH (308 nm) increased as the pressure was increased. There were no clear trends in the changes with repeat measurements. Because the probe was repositioned in the same location for repeat measurement, in addition to slight changes in position, there may also be slight differences in the pressure applied during a given measurement. The poor result for cHb is surprising, as it would be expected that pressure would have a greater effect. Clearly the probe contact can

affect the measurements, however, the results for the complete set of data capture the effects of these variations reasonably well.

We did not attempt to quantify the percent change in the parameters or signal intensity as these probe properties were varied because these variables were not studied in a rigorous manner. Changes in lineshape and intensity could be appreciated both in the reflectance and fluorescence measurements. In the case of varying the pressure, highly compressible sites (BM, VT, FM) tended to be more strongly affected than less compressible sites (i.e. DT, LT).

5.4 Conclusions

The assessment of the natural variation of healthy oral tissue is a crucial step in developing a spectroscopic-based tool for cancer diagnosis. Because the oral cavity is complex, in that it includes several different anatomic sites, the goal of this work was to characterize the natural variations. In this study we employed reflectance and intrinsic fluorescence spectroscopy to characterize 9 sites in the oral cavity with a physical-based model. In comparing the parameters among various sites, we have shown that there are significant differences in the mean and spread for various sites and therefore spectroscopy is influenced by anatomy. Our findings also demonstrate the strong correlation between the parameters and the physical properties of the tissue. Keratinized sites (HP, GI) were distinct from non-keratinized sites based on cluster analysis, as well as the model-based parameters. We have identified which parameters are affected by this feature and how the parameters are affected. Even

among non-keratinized sites, certain sites could be distinguished with >80% sensitivity and specificity with only 2 parameters. Results from the clustering give an indication of which sites are similar or dissimilar and display unique spectral properties. In addition, we have developed thresholds for the range of normal values for each parameter for each site.

Our understanding of the impact of keratin on the spectral parameters will be valuable in our assessment of patient data, as many lesions (benign, dysplastic, or malignant) exhibit hyperkeratosis (leukoplakia). Using the thresholds for the range of normal values for each parameter for each site, we have the potential to determine whether we can detect malignancy in the absence of visible findings (Chapter 6).

The findings in this work also highlight a number of major points to be considered in the application of spectroscopy to the detection of oral cancer. First, differences in the distribution of white (keratinized) lesions among benign, dysplastic, and cancerous data sets could be interpreted mistakenly as evidence of a separation between benign and malignant tissue. This is shown by the frequent separation between the HP and GI and other non-keratinized sites. This would have a particularly pronounced impact when data collected from healthy volunteers are directly compared to data collected from patients with lesions, as oral lesions are often white (leukoplakia). When small numbers of benign samples are grouped with large numbers of samples collected from healthy volunteers without visible lesions, this effect is also likely to occur.

Second, this work also emphasizes that by combining several anatomic sites, differences in anatomy can skew distributions, such that in comparing malignant

tissue samples to benign/healthy samples, each of which are comprised of different proportions of each site, a separation that is completely unrelated to malignancy may arise. We demonstrated cases in which sensitivities and specificities of ≥80% were achieved in separating different anatomic sites.

Finally, combining sites that have statistically different mean values can result in a larger spread in each parameter, and this makes it more difficult for a change that occurs with cancer to become significant. The combination of a site with a large spread with a site with a smaller spread would also make it more difficult to appreciate changes with disease for the site with the smaller spread. These effects are clearly demonstrated by the group statistics for the combination of 9 sites as compared to a single site.

Because we have fully characterized how various sites compare to one another for the various parameters, for future studies of lesions we may be able to combine certain sites with similar properties. The number of biopsy cases demonstrating dysplasia or cancer is usually limited, and combining sites would increase the number of samples from which a diagnostic algorithm could be developed, and also extend the number of sites to which it could be applied. The disadvantage is that this may incur a decrease in sensitivity and specificity for cancer due to an increase in the overall spread for the distributions by combining sites.

This work provides an important foundation for guiding our analysis of the patient data and the development of spectroscopic-based diagnostic algorithms for oral cancer.

5.5 References

1. D. C. G. de Veld, M. Skurichina, M. J. H. Witjes, R. P. W. Duin, H. J. C. M. Sterenborg and J. L. N. Roodenburg, "Clinical study for classification of benign, dysplastic, and malignant oral lesions using autofluorescence spectroscopy," *Journal of Biomedical Optics*, **9**(5), 940-950 (2004).
2. G. S. Nayak, S. Kamath, K. M. Pai, A. Sarkar, S. Ray, J. Kurien, L. D'Almeida, B. R. Krishnanand, C. Santhosh, V. B. Kartha and K. K. Mahato, "Principal component analysis and artificial neural network analysis of oral tissue fluorescence spectra: Classification of normal premalignant and malignant pathological conditions," *Biopolymers*, **82**(2), 152-166 (2006).
3. M. G. Muller, T. A. Valdez, I. Georgakoudi, V. Backman, C. Fuentes, S. Kabani, N. Laver, Z. M. Wang, C. W. Boone, R. R. Dasari, S. M. Shapshay and M. S. Feld, "Spectroscopic detection and evaluation of morphologic and biochemical changes in early human oral carcinoma," *Cancer*, **97**(7), 1681-1692 (2003).
4. N. Subhash, J. R. Mallia, S. S. Thomas, A. Mathews and P. Sebastian, "Oral cancer detection using diffuse reflectance spectral ratio R540/R575 of oxygenated hemoglobin bands," *Journal of Biomedical Optics*, **11**(1), (2006).
5. D. L. Heintzelman, U. Utzinger, H. Fuchs, A. Zuluaga, K. Gossage, A. M. Gillenwater, R. Jacob, B. Kemp and R. R. Richards-Kortum, "Optimal excitation wavelengths for in vivo detection of oral neoplasia using fluorescence spectroscopy," *Photochemistry and Photobiology*, **72**(1), 103-113 (2000).
6. A. Sharwani, W. Jerjes, V. Salih, B. Swinson, I. J. Bigio, M. El-Maaytah and C. Hopper, "Assessment of oral premalignancy using elastic scattering spectroscopy," *Oral Oncology*, **42**(4), 343-349 (2006).
7. H. J. van Staveren, R. L. P. van Veen, O. C. Speelman, M. J. H. Witjes, W. M. Star and J. L. N. Roodenburg, "Classification of clinical autofluorescence spectra of oral leukoplakia using an artificial neural network: a pilot study," *Oral Oncology*, **36**(3), 286-293 (2000).
8. H. M. Chen, C. P. Chiang, C. You, T. C. Hsiao and C. Y. Wang, "Time-resolved autofluorescence spectroscopy for classifying normal and premalignant oral tissues," *Lasers in Surgery and Medicine*, **37**(1), 37-45 (2005).
9. V. R. Kolli, H. E. Savage, T. J. Yao and S. P. Schantz, "Native Cellular Fluorescence Of Neoplastic Upper Aerodigestive Mucosa," *Archives of Otolaryngology-Head & Neck Surgery*, **121**(11), 1287-1292 (1995).
10. A. Gillenwater, R. Jacob, R. Ganeshappa, B. Kemp, A. K. El-Naggar, J. L. Palmer, G. Clayman, M. F. Mitchell and R. Richards-Kortum, "Noninvasive diagnosis of oral neoplasia based on fluorescence spectroscopy and native tissue autofluorescence," *Archives of Otolaryngology-Head & Neck Surgery*, **124**(11), 1251-1258 (1998).
11. J. K. Dhingra, D. F. Perrault, K. McMillan, E. E. Rebeiz, S. Kabani, R. Manoharan, I. Itzkan, M. S. Feld and S. M. Shapshay, "Early diagnosis of

upper aerodigestive tract cancer by autofluorescence," *Archives of Otolaryngology-Head & Neck Surgery*, **122**(11), 1181-1186 (1996).

12. D. C. G. de Veld, M. Skurichina, M. J. H. Witjes, R. P. W. Duin, D. J. C. M. Sterenborg, W. M. Star and J. L. N. Roodenburg, "Autofluorescence characteristics of healthy oral mucosa at different anatomical sites," *Lasers in Surgery and Medicine*, **32**(5), 367-376 (2003).
13. M. G. Muller, I. Georgakoudi, Q. G. Zhang, J. Wu and M. S. Feld, "Intrinsic fluorescence spectroscopy in turbid media: disentangling effects of scattering and absorption," *Applied Optics*, **40**(25), 4633-4646 (2001).
14. R. R. Wilcox, *Fundamentals of modern statistical methods : substantially improving power and accuracy*, Springer, New York, 2001
15. I. Paetau, D. Rao, E. R. Wiley, E. D. Brown and B. A. Clevidence, "Carotenoids in human buccal mucosa cells after 4 wk of supplementation with tomato juice or lycopene supplements," *American Journal of Clinical Nutrition*, **70**(4), 490-494 (1999).
16. Y. S. Peng, Y. M. Peng, D. L. McGee and D. S. Alberts, "Carotenoids, Tocopherols, And Retinoids In Human Buccal Mucosal Cells - Intraindividual And Interindividual Variability And Storage Stability," *American Journal of Clinical Nutrition*, **59**(3), 636-643 (1994).
17. R. Brandt, G. E. Kaugars, W. T. Riley, Q. Y. Dao, S. Silverman, J. G. L. Lovas and B. P. Dezzutti, "Evaluation Of Serum And Tissue-Levels Of Beta-Carotene," *Biochemical Medicine and Metabolic Biology*, **51**(1), 55-60 (1994).
18. H. Du, R. C. A. Fuh, J. Z. Li, L. A. Corkan and J. S. Lindsey, "PhotochemCAD: A computer-aided design and research tool in photochemistry," *Photochemistry and Photobiology*, **68**(2), 141-142 (1998).
19. A. Amelink, H. J. C. M. Sterenborg, M. P. L. Bard and S. A. Burgers, "In vivo measurement of the local optical properties of tissue by use of differential path-length spectroscopy," *Optics Letters*, **29**(10), 1087-1089 (2004).
20. R. L. P. van Veen, W. Verkruysse and H. J. C. M. Sterenborg, "Diffuse-reflectance spectroscopy from 500 to 1060 nm by correction for inhomogeneously distributed absorbers," *Optics Letters*, **27**(4), 246-248 (2002).
21. T. A. Winning and G. C. Townsend, "Oral mucosal embryology and histology," *Clinics in Dermatology*, **18**(5), 499-511 (2000).
22. S. S. Sternberg, *Histology for pathologists*, Lippincott Williams & Wilkins, New York, 1997
23. A. R. Ten Cate, *Oral histology : development, structure, and function*, Mosby, St. Louis, 1998
24. S. K. Chang, N. Marin, M. Follen and R. Richards-Kortum, "Model-based analysis of clinical fluorescence spectroscopy for in vivo detection of cervical intraepithelial dysplasia," **11**(2), (2006).
25. T. Collier, M. Follen, A. Malpica and R. Richards-Kortum, "Sources of scattering in cervical tissue: determination of the scattering coefficient by confocal microscopy," *Applied Optics*, **44**(11), 2072-2081 (2005).

26. Y. C. Wu and J. N. Y. Qu, "Autofluorescence spectroscopy of epithelial tissues," *Journal of Biomedical Optics*, **11**(5), - (2006).
27. E. J. Reith and M. H. Ross, *Atlas of descriptive histology*, Medical Dept., New York,, 1970

CHAPTER 6

In Vivo Clinical Study: Patients

In this chapter we describe the development of spectral diagnostic algorithms using data collected from patients with lesions. We first use the healthy volunteer data to identify spectrally similar anatomic sites. Next, we develop diagnostic algorithms to distinguish healthy mucosa from lesions or subsets of lesions. Finally, we develop spectral algorithms to distinguish benign lesions from dysplastic/malignant lesions.

6.1 Motivation and Goals

Biopsy, the current gold standard for the detection of oral cancer, is limited by the fact that it is invasive, and the final diagnosis is based on a qualitative assessment of the tissue morphology[1]. In contrast, optical spectroscopy can provide quantitative information about tissue morphology and biochemistry, and may therefore non-invasively reveal changes in the properties of the tissue as it is altered by the disease process. A spectroscopic tool could potentially enable the physician to assess a lesion area over time without repeated biopsies, confirm the need for biopsy and the most appropriate area to be biopsied for a patient with a suspicious lesion, and permit several regions of a lesion to be evaluated beyond what may be typically biopsied.

The goal of the patient study is to develop spectroscopic diagnostic algorithms with which we can distinguish healthy mucosa from lesions (clinically abnormal mucosa), and benign lesions from dysplastic/malignant lesions. Four types of comparisons were performed in our analysis: 1) Healthy mucosa versus benign lesions, 2) Healthy mucosa versus dysplastic/malignant lesions, 3) Healthy mucosa versus all lesions (benign, dysplastic, malignant), and 4) Benign lesions versus dysplastic/malignant lesions.

The distinction of healthy mucosa from lesions is not of major clinical importance; however, we take advantage of the relatively large numbers of samples within each category to test whether certain anatomic sites which are spectrally similar can be combined and preserve or even improve diagnostic performance as compared to when all sites are combined. Because differences between healthy

mucosa and lesions are visible by eye, we would expect excellent performance for the diagnostic algorithms. Distinguishing dysplastic/malignant lesions from benign lesions has considerable clinical significance, as the former group is more likely to result in intervention. For this separation, we also compare the results for the case in which all anatomic sites were combined, a single anatomic site was examined, or spectrally similar anatomic sites were combined. For each comparison, we also examined the trends in the extracted physical parameters with the development of dysplasia/malignancy.

6.2 Materials and Methods

6.2.1 In Vivo Data Collection

Subjects were recruited from Boston Medical Center (BMC). The protocol for *in vivo* data collection was approved by the institutional review board at BMC and the Committee on the Use of Humans as Experimental Subjects at the Massachusetts Institute of Technology. Subjects enrolled in the study included patients with suspicious lesions undergoing biopsy in the clinic or operating room, and patients undergoing surgical resection of known lesions. Written informed consent was obtained from all participants to indicate their willingness to participate in the study. Relevant background information was obtained from each patient, including their smoking and alcohol consumption, history of oral or upper aerodigestive tract lesions, and if relevant, any treatment received. Following collection of the spectral data the physician used a small punch biopsy (1.5 or 2 mm in diameter) to score the exact area

in which the probe was placed. The tissue sample was subsequently removed (biopsy) with a larger punch biopsy (3.5 mm diameter), and sent for pathological evaluation. In some cases a scalpel biopsy was performed because of the lesion architecture or location. A microscopic assessment of each sample was performed by 3 pathologists and the consensus diagnosis was considered to be the final diagnosis. The samples were classified as either benign, dysplastic, malignant, or indefinite.

The set of 710 spectra collected from the healthy volunteers (HV) (Chapter 5) were additionally used in the present study, as will be described, to represent healthy mucosa.

6.2.2 Model-based Analysis

The spectra were calibrated and the modeled fit obtained as detailed in Chapters 3.3 and 4.2, in the same manner as for the HV data. The diffuse reflectance spectra were fit over the range of 380-700 nm. From the reflectance spectra we derived 7 parameters: A, B, C, concentration of hemoglobin (cHb), the oxygen saturation (α), effective vessel radius, and concentration of β-carotene ($c\beta C$). Because the results from the HV study showed that the area-normalization of the fluorescence spectra yielded parameter distributions with smaller relative standard deviations, all fluorescence parameters were extracted following this pre-processing step (Section 5.2.2). Analysis of the area-normalized intrinsic fluorescence, as before, yielded 6 parameters: Tryptophan (308 nm), Collagen (308 nm), NADH (308 nm), NADH (340 nm), Coll401 (340 nm), and Coll427 (340 nm). The number in parentheses indicates the excitation wavelength. In order to be able to apply the findings from

health volunteer study to the patient data, the 308 nm and 340 nm fluorescence excitations were used for this analysis.

6.2.3 Identifying Spectrally Similar Anatomic Sites

Because of the limited numbers of samples in each lesion category (benign, dysplastic, malignant, or indefinite) for a given anatomic site, we used the HV data to determine if data collected from certain anatomic sites could be combined. We used two methods to test whether specific sites were similar enough in their spectral profile (i.e. parameter distributions), such that they could be combined into a single group and anatomy would not influence the performance of the spectral algorithms. Both methods were applied to the HV data after the interquartile range exclusion criteria had been applied. The first method was a multiple comparison test of the mean values for each parameter for each site. Using this method, the mean values for two sites were considered statistically different if the p-value was <0.05. The Tukey-Kramer correction was applied to compensate for multiple comparisons. The second method used to compare the distributions for the sites was the Kruskal-Wallis test, a generalization of the Wilcoxon rank sum test to allow greater than two samples to be compared. This test is a nonparametric alternative to ANOVA and in contrast to the previous method, does not assume that the parameters follow a specific type of distribution (i.e. normal distribution)[2]. Using this method we test whether the medians of the distributions for two samples are identical. Following this step, a multiple comparison test was performed of the ranks for the groups using the Tukey-Kramer correction. Both tests were performed at the 5% significance level.

For each method we calculated a "similarity score" which represents the total number of instances two sites were found not to be significantly different when their distributions for each of the 13 parameters were compared. Perfect correspondence between the parameters for two sites would be indicated by a similarity score of 13.

6.2.4 Spectral Diagnostic Algorithm Development

The details of the statistical methods used in the process of developing the diagnostic algorithms are described in Chapter 4.3. Diagnostic algorithms were developed for distinguishing specific subsets of data using logistic regression. Spectral parameters with the greatest diagnostic potential for the comparison of interest were identified using the log likelihood ratio (LLR) test. The fewest parameters necessary to achieve the maximum ROC-AUC for a comparison were retained. To prevent overtraining the data, we imposed an additional criterion on the maximum number of parameters that could be included in the model. At least 5 training samples per class were required for each spectral parameter included in the model. For example, for a comparison of 13 healthy mucosa samples versus 25 lesion samples, the minimum number of samples in a class is 13 therefore a maximum of 2 diagnostic parameters were used in the model. Ideally, the ratio should be at least 10 parameters per training samples in the class, however this was not practical given the limited number of samples in our data[3]. The discriminatory power of each algorithm was evaluated by performing leave-one-out cross-validation (LCV) and calculating the area under the receiver operator characteristic curve (ROC-AUC). The sensitivity and specificity were additionally determined.

6.3 Results and Discussion

6.3.1 General Description of the Data

A total of 79 subjects were recruited in the study. Of the 114 spectra collected from these subjects, 13 spectra had to be excluded due to an error noted during data collection, a broken probe, the poor quality of the biopsy specimen, or the pathology being unavailable. Data were collected from the following sites in the oral cavity: buccal mucosa (BM), dorsal surface of the tongue (DT), floor of the mouth (FM), gingiva (GI), hard palate (HP), lateral surface of the tongue (LT), retromolar trigone (RT), soft palate (SP), and ventral surface of the tongue (VT), tonsils (TL), base of tongue (BT), lips (IL), and larynx (LX). Because we did not have comparable data from healthy volunteers, the 12 spectra collected from the TL, BT, IL, and LX were excluded from the analysis. Two additional samples, a lesion which turned out to be a lymphoma, and a sample classified as indefinite, were also excluded. A total of 87 spectra collected from 67 subjects were included in the analysis. A summary of the number of spectra measured from the subjects by disease category and site is shown in Table 6.2. The frequency of lesions among the various sites in our data is consistent with the well known high risk sites for oral cancer, which include the tongue and floor of the mouth[4]. The average age (± standard deviation) for the subjects was 56±16 years. The HV data used in the analysis in this chapter are summarized in Table 5.1. Specific comparisons involving a single anatomic site were made in the subsequent analysis for the BM, LT, FM, SP, and VT. These groups consisted of 100, 112, 112, 43, and 74 spectra, respectively.

Site	Malignant	Dysplastic	Benign	Indefinite	Total
Buccal mucosa	2	0	11	1	**14**
Dorsal tongue	0	0	4	0	**4**
Floor of mouth	2	11	3	0	**16**
Gingiva	0	1	0	0	**1**
Hard palate	1	1	2	0	**4**
Lateral tongue	4	9	12	0	**25**
Retromolar trigone	0	1	3	0	**4**
Soft palate	4	0	5	0	**9**
Ventral tongue	1	3	7	0	**11**
Total	**14**	**26**	**47**	**1**	**88**

Table 6.1 Summary of the number of spectra measured from each anatomic site and the consensus pathological classification.

6.3.2 Identifying Spectrally Similar Anatomic Sites

From the results of the HV study, we found that anatomy strongly influences the spectral parameters. However, if we identify which anatomic sites are spectrally similar, lesions from these sites may be able to be combined without compromising diagnostic performance. The results for the similarity score (the total number of parameters for which two sites were not statistically different) based on the multiple comparison test and Kruskal-Wallis test are shown in Tables 6.2 an 6.3, respectively.

Site	BM	DT	FM	GI	HP	LT	RT	SP	VT
BM		4	5	3	1	4	8	9	5
DT			6	6	5	2	9	6	2
FM				3	1	6	3	8	10
GI					11	3	6	7	2
HP						2	6	4	0
LT							4	4	8
RT								7	3
SP									7
VT									

Table 6.2 Table listing the total number of instances in which the mean values of a parameter for a pair of sites were characterized as statistically similar (not significantly different), according to a multiple comparison test performed for each of the 13 parameters. Perfect correspondence between two sites would yield a score of 13.

Site	BM	DT	FM	GI	HP	LT	RT	SP	VT
BM		4	5	2	1	5	8	10	5
DT			6	6	6	3	9	7	3
FM				3	1	5	3	9	12
GI					11	3	7	7	2
HP						3	6	4	0
LT							5	6	8
RT								8	3
SP									8
VT									

Table 6.3 Table listing the total number of instances in which the median values of a parameter for a pair of sites were characterized as statistically similar (not significantly different), according to a Kruskal-Wallis test performed for each of the 13 parameters. Perfect correspondence between two sites would yield a score of 13.

The scores and trends are fairly similar for both methods. Both tests demonstrate that the various anatomic sites demonstrate pronounced spectral differences. For the multiple comparison test, 26 of the 36 comparisons (72%) produced a similarity score of <7. Similarly, 24 of the 36 comparisons (67%) for the

Kruskal-Wallis test yielded scores <7. For each site, we identified which of the remaining sites with which it demonstrated the highest score. We identified 4 groups of sites with similar spectral properties (albeit to different degrees) which were common to both methods: 1) BM and SP, 2) FM and VT, 3) DT and RT, and 4) GI and HP.

In order to benefit from the ability to combine spectrally similar sites and improve our ability to create spectral diagnostic algorithms despite limited numbers of lesion samples, there are two important considerations. First, the total number of samples for the combined set sites must be sufficient for further analysis. Second, the number of samples in each category to be compared must be sufficient. In our case the categories of particular interest are benign lesions versus dysplastic/malignant lesions. Table 6.4 lists the number of samples in each of the lesion categories for the spectrally similar sites. Two of the four groups of spectrally similar sites met these criteria. Even by combining the GI and HP or DT and RT, the number of samples was not substantially increased and therefore we did not perform any of the comparisons described in the subsequent sections for these specific groups.

Similarity Score Multiple Comparison test/ Kruskal-Wallis test	Spectrally Similar Sites	Malignant	Dysplastic	Total	Benign	Total
9 / 10	Buccal mucosa	2	0	6	11	16
	Soft palate	4	0		5	
10 / 12	Floor of the mouth	2	11	17	3	10
	Ventral surface of the tongue	1	3		7	
9 / 9	Dorsal surface of the tongue	0	0	1	4	8
	Retromolar trigone	0	1		4	
11 / 6	Gingiva	0	1	3	0	2
	Hard palate	1	1		2	

Table 6.4 Spectrally similar anatomic sites as determined by the consensus of the results for the multiple comparison test and Kruskal-Wallis test. The number of dysplastic/malignant, and benign sample subtotals and totals are shown for each group.

6.3.3 Distinguishing Healthy Mucosa from Lesions, Subsets of Lesions

In Tables 6.5 and 6.6 we compare clinically normal mucosa (from healthy volunteers) to clinically abnormal mucosa (from patients). For each of the three comparisons listed, the ROC-AUC value for the combination of all 9 sites, the combination of all sites except the GI and HP (combination of 7 sites), individual sites, and spectrally similar sites are shown. We considered the combination of all sites except the GI and HP (combination of 7 sites) because the k-means clustering results from the HV study (Chapter 5) clearly demonstrate that the GI and HP were spectrally distinct from most other sites for a number of parameters. For comparison, we also list the results for the combination of 2 spectrally dissimilar anatomic sites, LT/BM (similarity scores of 4 and 5 for the multiple comparison test and Kruskal-Wallis test, respectively).

The combination of the 9 sites yielded 710 healthy mucosa samples, 47 benign samples, and 40 dysplastic/malignant samples. The combined group of 7 sites

consisted of 598 healthy mucosa samples, 45 benign lesions, and 37 dysplastic/malignant samples. For some comparisons, the number of samples in one or both categories was too few to be analyzed reliably.

Group	Healthy Mucosa vs. Lesions	Healthy Mucosa vs. Dysplastic/Malignant Lesions	Healthy Mucosa vs. Benign Lesions
	Maximum ROC-AUC		
9 Sites Combined	0.83	0.85	0.81
7 Sites Combined	0.84	0.88	0.84
LT	0.95	0.97	0.95
FM	0.94	0.95	---
BM*	0.94	---	0.96
VT	0.95	---	---
FM/VT	0.97	0.97	0.94
BM/SP*	0.91	---	0.90
LT/BM	0.92	---	0.91

Table 6.5 Summary of the maximum ROC-AUC values obtained for three different separations for the combination of all sites, individual sites, and the combination of spectrally similar sites. *No dysplastic samples.

The diagnostic performance of the spectral algorithms was poorest for the group in which all 9 sites were combined. Only a slight increase in the ROC-AUC was produced by omitting these sites as compared to the combination of 9 sites, however, the combined number of GI and HP lesion samples was very small (5 samples total). A similar finding was observed in a study by Muller *et al.*, in which they examined the diagnostic performance of their spectral diagnostic algorithms for distinguishing lesions from healthy/benign tissue, and dysplasia from malignant tissue. They found that the best sensitivity and specificity was obtained when samples from 4 non-keratinized anatomic sites (BM, FM, SP, VT) were analyzed separately from keratinized anatomic sites (DT, GI, HP). Furthermore, when they combined these two groups the overall sensitivity and specificity for the separations decreased.

When we examined the same 4 keratinized sites, 4 parameters were found to be statistically significant for separating healthy mucosa from lesions (in order of decreasing significance): cβc, Coll308, cHb, and Tryptophan (308 nm). For the first two parameters all 4 sites are statistically similar according to both the multiple comparison and Kruskal-Wallis tests. An ROC-AUC value of 0.92 was obtained.

Most investigations examining the potential of spectroscopy to non-invasively identify oral cancer attempt to distinguish healthy mucosa from visible lesions (dysplastic/malignant lesions or malignant lesions only). Benign lesions are frequently excluded from the analysis for reasons which are not specified[5-8]. Approximately 80% of leukoplakias, the most frequently encountered lesion of the oral cavity, prove to be benign upon pathological evaluation[9]. Therefore, it is important to consider this lesion category. The distinction between clinically abnormal and clinically normal tissue is easily appreciated by eye and it would be expected that spectroscopy would also easily appreciate the differences between these two groups. The results in Table 6.5 indicate that excellent separation between these two groups is possible using spectroscopy, particularly in the cases where the diagnostic algorithms are developed and applied to individual sites, or spectrally similar sites. The excellent separations achieved for the three comparisons are likely due to the influence of keratinization, ulceration, and other abnormal mucosal changes associated with the lesions, which are not directly linked to malignancy. This is supported by the fact that the comparison of healthy mucosa versus benign lesions yields both similar ROC-AUC values and diagnostic parameters as for the comparison against dysplastic/malignant lesions alone. The decreased accuracy of the

algorithms when applied to the combination of 9 is most likely due to the influence of anatomic differences. Although many studies obtain excellent results while combining all sites, usually the total number of samples (or patients) is small and/or the majority of the lesion samples are derived from a few anatomic sites[10-12]. Comparisons of healthy mucosa versus subsets of lesions may be particularly affected by anatomic differences, as some sites are more represented than others in the different lesion categories. As shown in Table 6.1, almost all the BM lesions are benign, while the majority of the FM lesions are dysplastic/malignant. These differences could falsely increase the diagnostic accuracy, if the sites in the two categories to be compared are spectrally distinct.

Figure 6.1 shows examples of the ROC curves obtained for the separation of healthy mucosa versus all lesions for the combination of 9 sites, LT, and FM/VT.

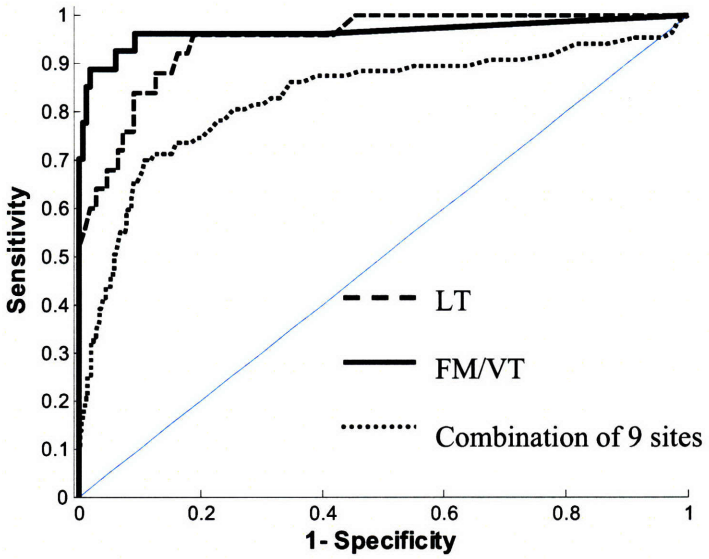

Figure 6.1. Examples of ROC curves obtained for the separation of healthy mucosa from lesions for the combination of nine sites (dotted line), LT (dashed line), and the FM/VT (solid line). The thin blue line represents the ROC curve that would result due to random chance.

The diagnostic parameters which were combined to yield the ROC-AUC values shown in Table 6.5 are shown in Table 6.6. The parameters are listed in order of decreasing significance based on the results of the log likelihood ratio test. As shown in the table, 2-4 parameters were found to be significant between the two categories for all the comparisons shown in Table 6.6.

For the comparison of healthy mucosa versus all lesions, 6 parameters were frequently noted as being significant as shown in Table 6.6: cβC, Tryptophan (308 nm), NADH (340 nm), cHb, Coll427 (340 nm), and α. The effective vessel radius, A and C parameters were never among the diagnostic parameters. For the LT, the B parameter was most significant, a parameter which was not significant for any other separation. NADH (308 nm) was important for the BM as an individual site. In examining the combination of spectrally dissimilar sites, LT/BM, NADH (308 nm) and cβC, were both parameters for which the two sites were spectrally similar, which may explain the high ROC-AUC values.

	Healthy Mucosa vs. Lesions	**Healthy Mucosa vs. Dysplastic/Malignant Lesions**	**Healthy Mucosa vs. Benign Lesions**
Group	**Diagnostic Parameters**		
9 Sites Combined	NADH340, cHb, α	NADH340, cHb	Coll427, cHb
7 Sites Combined	NADH340, α, cHb	NADH340, cβC, cHb	NADH340, Trp308, α
LT	B, Trp308, α	Coll401, B,Trp308	B, Trp308, NADH340
FM	Trp308, cβC	Trp308, Coll427	---
BM*	NADH340, NADH308	---	NADH340, NADH308
VT	Coll427, cβC	---	---
FM/VT	Coll427, Trp308, cβC	Coll427, α	Coll427, cβC
BM/SP*	cHb, cβC, NADH340	---	NADH340, cβC
LT/BM	NADH340, NADH308 cβC	---	NADH340, NADH308, cβC

Table 6.6 Summary of the significant diagnostic parameters determined for three comparisons. The parameters are written in order of diagnostic significance. Trp308 (Tryptophan (308 nm)), Coll308 (Collagen (308 nm)), NADH308 (NADH (308 nm)), NADH340 (NADH (340 nm)), Coll401 (Coll401 (340 nm)), Coll427 (Coll427 (340 nm)) *No dysplastic samples.

For all the comparisons, the order of significance of these parameters is also relatively consistent. Therefore, the difference in the diagnostic accuracy for the combination of 7 and 9 sites versus the individual and spectrally similar groups is most likely related to the fact that the threshold for the combined sites is not as optimal. For the Coll427 (340 nm) contribution, the combination of 9 sites, VT and FM/VT groups had similar relative standard deviations (RSDs) for their respective healthy mucosa samples and lesion samples (range 13-20% and 36-41%, respectively). The upper limits of the ranges were for the combined group. However, the ratios of the means for the healthy mucosa compared to the lesions for the three groups were considerably different: 1.29, 1.63, and 1.62, respectively. This suggests that in combining numerous sites spectral contrast is diminished.

In Table 6.7, we list the trends in the parameters for lesions as compared to healthy mucosa and compare the trends in the parameters for the GI and HP (keratinized sites) relative to that of the other anatomic sites. We use the latter comparison as a surrogate marker for the impact of keratin, although this is not a perfect comparison because other mucosal features specific to these 2 sites may contribute to the trends. All the trends in the parameters for the GI and HP relative to the other anatomic sites were statistically significant except for Coll427 (340 nm). The trends in each parameter when lesions or subsets of lesions were compared to healthy mucosa were found to be independent of the specific comparison (i.e. versus lesions, dysplastic/malignant lesions, or benign lesions).

Group	GI/HP	Lesions
A	Decrease	---
B	Increase	Increase
C	Decrease	---
cHb	Decrease	Increase
α	Increase	Increase
Effective vessel radius	Decrease	---
cβC	Decrease	Decrease
Tryptophan (308 nm)	Increase	Increase
Collagen (308 nm)	Decrease	---
NADH (308 nm)	Decrease	Decrease
NADH (340 nm)	Increase	Increase
Coll401 (340 nm)	Decrease	Decrease
Coll427 (340 nm)	NSD	Decrease

Table 6.7 Changes in the parameters that occur with the presence of keratin (as determined by a comparison of the GI/HP to the other 7 anatomic sites) and for visible lesions relative to healthy mucosa. NSD: not significantly different

As is shown in the table, the trends in the parameters for both groups are identical for 7 of the 9 parameters. The trends were reversed for cHb. The Coll427 (340 nm) contribution decreased for the lesions compared to healthy mucosa, but was not statistically different for the keratinized sites as compared to the other sites. The strong parallel in the trends for the GI/HP and lesions provide added support that keratin may be a key contributing factor in the differences between lesions and healthy mucosa.

A study by Amelink *et al.* investigated the use of differential path length spectroscopy to distinguish healthy mucosa from malignant lesions using a model-based approach that incorporated vessel packaging[13]. They also found that malignant lesions demonstrated a significant decrease in the hemoglobin oxygenation, increase in the blood content, and increase in the scattering slope (B parameter). A significant decrease in scattering amplitude (A parameter) was also observed, which was not found in our analysis. A study by Muller *et al.* similarly found that for fluorescence excited at 340 nm, the contribution of NADH increased and collagen decreased with

disease progression[10]. Muller *et al.* combined intrinsic fluorescence spectroscopy and diffuse reflectance spectroscopy to evaluate the oral lesions.

The increase in the B parameter is most likely due to the scattering from keratin present on leukoplakias. In addition to the impact of keratin, the increased cHb may be due to angiogenesis or the vascular dilation seen in erythroplakias (Section 2.2,2). The decreased collagen is consistent with degradation of the extracellular matrix by matrix metalloproteinases (Section 3.1). The increase in NADH at 340 nm excitation may result from increased metabolic activity.

6.3.4 *Distinguishing Benign Lesions from Dysplastic/Malignant Lesions*

In order to determine whether the reflectance and fluorescence parameters can be used to identify the most severe lesions, we developed diagnostic algorithms to be applied to clinically abnormal lesions. This distinction has considerable clinical significance, as the differences in the two categories cannot be appreciated by observation alone. Spectral algorithms designed to distinguish benign lesions from dysplastic/malignant lesions were developed for the 5 groups with sufficient numbers of samples in the two lesion categories: 9 sites combined, 7 sites combined, LT, FM/VT, and BM/SP.

The ROC-AUC value, sensitivity, and specificity for the comparisons between benign lesions and dysplastic/malignant lesions are shown in Table 6.8. Unlike the comparisons involving healthy mucosa, the diagnostic performance is considerably poorer. The highest ROC-AUC value was obtained for the LT and the lowest value

for the combination of 9 sites. While the diagnostic performance for the FM/VT group is considerably better than that of the combination of 7 sites or 9 sites, the results for the BM/SP group is comparable.

Group	Benign Lesions vs. Dysplastic/Malignant Lesions Maximum ROC-AUC	Sensitivity/ Specificity [%]
9 Sites Combined	0.60	55.0 / 61.7
7 Sites Combined	0.63	56.8 / 64.4
LT	0.75	92.3 / 66.7
FM/VT	0.71	94.1 / 60.0
BM/SP*	0.63	83.3 / 56.2

Table 6.8 Summary of the maximum ROC-AUC values obtained for the separation of benign lesions from dysplastic/malignant lesions. *No dysplastic samples.

Figure 6.2 shows examples of the ROC curves obtained for the separation of benign lesions from dysplastic/malignant lesions for the combination of 9 sites, LT, FM/VT, and the BM/SP.

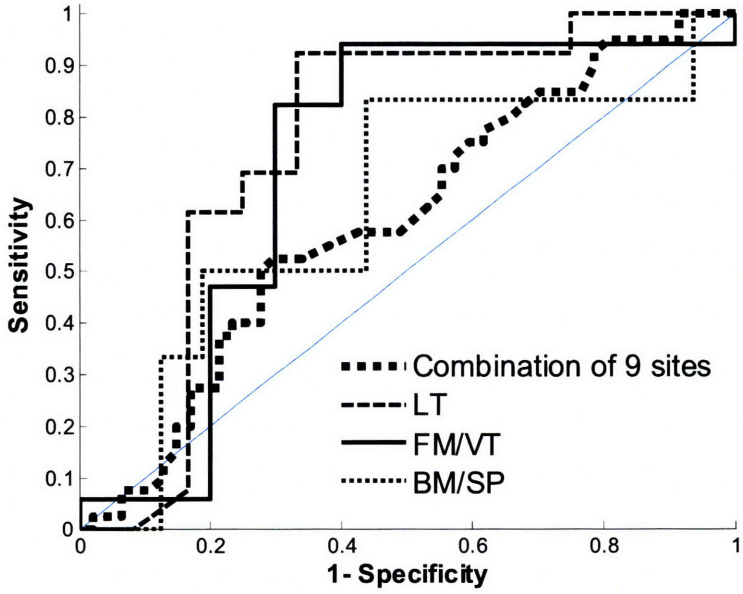

Figure 6.2. Examples of ROC curves obtained for the separation of benign lesions from dysplastic/malignant lesions for the combination of 9 sites (thick dashed line), LT (dashed line), FM/VT (solid line), and the BM/SP (dotted line). The thin blue line represents the ROC curve that would result due to random chance.

The diagnostic parameters that were combined to yield the ROC-AUC values shown in Table 6.8 are listed in Table 6.9.

Group	Benign Lesions vs. Dysplastic/Malignant Lesions Diagnostic Parameters
9 Sites Combined	C, Coll401, cβC
7 Sites Combined	C, cβC, Coll401
LT	cβC, Coll401
FM/VT	Coll427, C, cHb
BM/SP*	A

Table 6.9 Summary of the significant diagnostic parameters for separating benign lesions from dysplastic/malignant lesions. *No dysplastic samples. Coll401 (Coll401 (340 nm)), Coll427 (Coll427 (340 nm))

The important parameters for this comparison are almost entirely different than those which were important for the separation between healthy mucosa and lesions. While the C parameter was never a diagnostic parameter for distinguishing healthy mucosa from lesions or subsets of lesions, it played a significant role in separating benign lesions versus dysplastic/malignant lesions for 3 of the 5 groups. This was also the case for the Coll401 (340 nm) parameter. This parameter was important for 3 of the 5 groups for identifying dysplastic/malignant lesions, while only important in 2 instances for the comparisons against healthy mucosa. Just as for the comparisons against healthy mucosa, the cβC continued to be a diagnostic parameter. The consistency of the diagnostic performance among 4 of the 5 groups suggests that the parameters are not being influenced by anatomy, but intrinsic differences between the two lesion categories. In the case of the BM/SP group, most of the benign cases are from the BM and the malignant cases from the SP (there are no dysplastic lesions for this group), therefore anatomic differences likely play a role in the separation. This is supported by the fact that the HV data shows that the two

sites are significantly different from one another in their distributions for the A parameter.

The dysplastic/malignant lesion sites are characterized by an increased C parameter and cHb, and decreased cβC, Coll401 and Coll427. This trend in the C parameter is consistent with increased scattering by small particles as may occur with increased epithelial proliferation. The increased cHb may be the result of vascular dilation and inflammation as seen in erythroplakias or with ulceration, features which are more likely to be associated with dysplasia/malignancy than white lesions (Section 2.2.2). The decreased collagen fluorescence may reflect disruption of the basement membrane by the invading malignant cells and decreased stromal collagen as a result of the action of matrix metalloproteinases.

The decrease in the cβC is more difficult to explain. β-carotene, which is found at high concentrations in orange vegetables, is an antioxidant (protects against free radical damage). β-carotene can also be metabolized to vitamin A, which is involved in the differentiation of normal epithelial cells[14]. Supplementation with β-carotene has been explored as a chemopreventive strategy in a number of studies of oral cancer (and other cancers), but shown limited or no success in large trials in humans[15,16]. Interestingly, a case-control study of 6832 men of Japanese ancestry found that over the 20 years in which they were followed, those who developed esophageal, laryngeal, and oropharyngeal cancer demonstrated significantly reduced β-carotene levels as compared to controls after adjusting for cigarette smoking and alcohol intake[17].

Because 3 parameters, C, cβC and Coll401 (340 nm) were found to be most significant for distinguishing the two classes of lesions, we examined whether there were certain groups of sites which were spectrally similar for all three of these parameters. These sites could potentially be combined without reducing the diagnostic performance, as seen for the combination of 9 sites. A single pair of sites was found to be similar for all three parameters based on the consensus results for the multiple comparison test and Kruskal-Wallis test: the FM and VT. According to the multiple comparison test, the GI and HP were similar for all 3 parameters, but only for 2 of the 3 parameters by the Kruskal-Wallis test. The BM and SP were identified as similar for all 3 parameters by the Kruskal-Wallis test, but only for 2 parameters by the multiple comparison test. We tested the performance of a diagnostic algorithm for distinguishing benign lesions from dysplastic/malignant lesions for the combination of the FM, LT, and VT. Both tests grouped these sites together for 2 of the three parameters (C, cβC). An ROC-AUC value of 0.61 was obtained using these parameters, which is worse than for the LT alone or the FM/VT combination. This is consistent with the overall low similarity score (based on all 13 parameters) obtained for the FM and LT.

Few investigators have attempted to distinguish benign lesions from dysplastic/malignant lesions, and in most cases several anatomic sites are combined. In a study by de Veld *et al.* using reflectance and fluorescence spectroscopy, they analyzed 115 spectra from 11 anatomic sites (the DT and lip were excluded) and used principal component analysis in combination with various normalization and classifier methods[18]. They report a maximum ROC-AUC value of 0.73 using the

combination of the two methods when they combined all 11 sites. A study by van Staveren *et al.* using fluorescence spectroscopy achieved a sensitivity and specificity of 31% and 67%, respectively, using an artificial neural network[12]. In the present study, when all 9 sites were combined a sensitivity and specificity of 55% and 62%, respectively, was achieved.

Consistent with our findings, greater success has been achieved for classifying lesions when the sites included in the analysis is limited. Wang *et al.* performed an *in vivo* fluorescence study of patients with lesions located only on the buccal mucosa[19]. They were able to obtain a sensitivity and specificity of 81% and 96%, respectively, for distinguishing benign from dysplastic/malignant lesions. We obtained a sensitivity and specificity of 92.3% and 66.7% for the LT site. A study by Mallia *et al.* investigated the use of fluorescence at 404 nm excitation to discriminate grades of oral cancer[20]. The authors found that they needed to exclude lesions samples from the vermilion border of the lip, dorsal and lateral surfaces of the tongue in order to develop successful diagnostic algorithms.

6.3.5 Spectroscopic Studies in the Literature

In Table 6.10 we provide a summary of fluorescence and reflectance studies in the literature which have been performed *in vivo* in humans. We list our own results from two of the common comparisons in the literature, healthy mucosa versus malignant lesions, and healthy mucosa versus dysplastic/malignant lesions, as well as for benign lesions versus dysplastic/malignant lesions. Several important considerations in these studies should be noted, in particular when comparing our

results to those of other investigators. First, many studies use visibly healthy mucosa in patients with lesions as the source of healthy mucosa. Second, many studies involve very few patients or lesions, and the lesions are derived from only a few anatomic sites. Only the study by de Veld *et al.* and our study involve more than 25 patients[18]. Inter-patient variability can contribute to the spread in the extracted parameters and potentially reduce diagnostic performance, whereas when numerous samples are taken from the same patient this effect is reduced. Third, as mentioned earlier, many studies exclude benign lesions. This may reduce the heterogeneity among the lesions and potentially increase the diagnostic performance in comparisons of healthy mucosa versus visible lesions. Fourth, many of the studies listed (e.g. Mallia *et al.*, Tsai *et al.*, and Gillenwater *et al.*) provide a sensitivity and specificity which were not calculated using cross-validation, but by applying a threshold to the data. Finally, most investigators compare clinically normal mucosa to visible lesions, which is not the most clinically relevant distinction.

Source Comparison	Sens./Spec. [%]	ROC-AUC	Diagnostic Method	Healthy Tissue: # Spectra (# people)	Lesions: # Spectra (# Pts.) Path.	Details
Present Work Nml vs. Ca Nml vs. Dys/Ca Nml vs. Be/Dys/Ca Be vs. Dys/Ca	 85.7 / 84.6 77.5 / 85.6 78.2 / 76.5 55.0 / 61.7	 0.88 0.85 0.83 0.60	DRS 380 -700 nm em IFS 308,340 nm ex 370-600 nm em	710 (79 HVs)	87 (67 Pts) 14 Ca 26 Dys 47 Be	Results are for 9 anatomic sites combined
Heintzelman, DL *et al.*, 2000 Nml/Be vs. Dys/Ca	 100 / 98	---	Fl Ratio 350,400 nm ex 472 nm em	**Calibration** 35 (9 HVs) 17 (11 Pts) **Validation** 274 (53 HVs) 4 (3 Pts)	**Calibration** 10 (9 Pts) 4 Dys 6 Ca **Validation** 2 (3 Pts) 2 Dys 2Ca	No benign sites
Wang CY *et al.*, 2003 Be vs. Dys/Ca	 81/96	---	Partial least squares, artificial neural network 330 nm ex 340-601 nm em	15 HVs	13 Ca	All lesions were from buccal mucosa
van Staveren HJ *et al.*, 2000 Nml vs. Be/Dys	 86 /100	---	Artificial neural network 420 nm ex 465-650 nm em	6 (2 HVs)	22(21 Pts) 9 Be 13 Dys	Only studied leukoplakias
Majumder SK *et al.*, 2000 Nml vs. Ca	 86 / 63	---	PCA + logistic regression 337 nm ex 350-700nm em	82 (25 Pts)	126 (25 Pts) 126 Ca	Limited lesions to those which were suspicious for SCC Did not perform cross-validation
Majumder SK *et al.*, 2005 Nml vs. Ca	 85-91/95-97 (RVM) 87-93/95-96 (SVM)	---	Relevance vector machines (RVM) Support vector machines (SVM) 337 nm ex 375-700 nm em	153 (13HVs) 88 (16 Pts)	83 (16 Pts) 83 Ca	

Source **Comparison**	**Sens./Spec. [%]**	**ROC-AUC**	**Diagnostic Method:** Fl. Ex Fl. Em	**Healthy Tissue:** # Spectra (# people)	**Lesions:** # Spectra (# Pts.) Path.	**Details**
Muller M *et al.*, 2003 Nml/Be vs. Dys/ Ca	91.2 / 96.4* all sites	---	NADH, collagen contribution + logistic regression	38 (15 HVs)	53 (15 Pts) 22 Be 19 Dys 12 Ca	Used TMS
	94.4 / 96.6* nonker sites		337, 358 nm ex 350-650 nm em			
Dys vs. Ca	67.9 / 85.7 all sites	---				
	72.7 / 90.0 nonker sites					
Gillenwater *et al.*, 1998 Nml vs. Lesions (unspecified)	82 / 100	---	Fluor. ratio 337 nm ex 635 nm em/(455-490 nm) em	95 (8 HVs)	33(12 Pts) after exclusions	Did not perform cross-validation
Tsai *et al.*, 2003 Be vs. Dys/Ca	81 / 87	---	330 nm ex 480/360 nm em ratio		104 20 Ca 27 Dys 57 Be	Only buccal mucosa lesions in study data Did not perform cross-validation
de Veld *et al.*, 2005			DRS, Fl, PCA, NN, multiple classifiers	581 (70 HVs)	115 (155 Pts.) 88 Be 11 Dys 16 Ca	Excluded sites where diagnosis clear 11 anatomic sites, excluded lip, DT
Nml vs. Ca	94 / 94	0.98				
Nml vs. Be/Dys/Ca	83 / 86	0.90	366-450 nm ex 455-867 nm em			
Be vs. Dys/Ca	---	<0.70				
Mallia *et al.*, 2008. Nml vs. Be (hyperplastic)	100/100	---	Fluor. ratio: 500/685nm em 404 nm ex	35 (35 HVs) **Calibration:** 4 Nml (4 Pts)	**Calibration:** 36 (20 Pts) 18 Ca 9 Dys 9 Be	Excluded vermillion border lip, LT, DT (algorithm did not work) Biopsies taken only for portion of measured area Did not perform cross-validation
Be vs. Dys	100/100	---	420-720 nm em	**Validation:** 4 Nml (4 Pts)	**Validation:** 21 (17 Pts) 6 Ca 5 Dys 7 Be	
Dys vs. Ca	100/100	---				

Table 6.10 Summary of *in vivo* spectroscopic studies of the oral cavity in which diagnostic performance was evaluated using either sensitivity, specificity, and/or ROC-AUC. Nml: normal, NN: neural network, Pts: patients, Be: benign, Dys: dysplastic, em: emission, ex: excitation HVs: healthy volunteers, PCA: principal component analysis

Further improvement in the diagnostic algorithms for distinguishing benign lesions from dysplastic/malignant lesions may be possible with a larger data set. For this comparison the log likelihood ratio test identified very few parameters as

statistically significant in the diagnostic model and higher p-values were obtained when comparing the means between spectral parameters for these two classes, thus indicating low spectral contrast. The heterogeneity of the lesions within each class also contributes to the challenge of separating the two groups. For example, some malignant lesions had small foci of invasion in a large area of relatively normal epithelium, while others demonstrated more diffuse disease throughout the entire specimen.

6.3.6 Application of Spectral Algorithms in Clinical Practice

Because of the low prevalence of oral cancer even among smokers, a spectroscopic screening tool would be beneficial if it demonstrated a reasonably high NPV (Section 4.3.3), and provided a quick, and easy means of isolating those at high risk for disease. The high NPV is important because most of the population which is screened will be negative for disease and the test should not frequently indicate the need for further evaluation unnecessarily. If a patient tested positive, they could then be referred to an otolaryngologist for further evaluation (a more thorough oral exam, biopsy, etc.)[21-23]. As another possible application, we can consider using the spectral diagnostic methods to distinguish benign lesions from dysplastic/malignant lesions in the setting of a patient presenting with a previously identified suspicious lesion.

If we consider leukoplakia only, (as the prevalence of erythroplakia is <1%), one study of 3256 leukoplakias found that 42.9% of biopsied FM lesions and 24.2% of tongue lesions showed dysplasia/malignancy[9]. The percentage of the 3256 cases of leukoplakia which were located at these sites was 6.8% and 8.6%, respectively. Given

that our most successful diagnostic algorithms represent a combination of the tongue and FM (LT and FM/VT), if we consider 1000 patients with leukoplakia of the tongue/FM, a total of 442 FM cases and 558 tongue cases would be expected based on the relative frequency of leukoplakia at these two sites. Of the 1000 cases, 32.5% would be expected to show dysplasia/malignancy (42.9% of FM lesions and 24.2% tongue lesions). Given that our sensitivities and specificities for the LT and FM/VT sites are similar, we use the average values of 93.2% and 63.3% as the sensitivity and specificity, respectively, of our spectral diagnostic test. Based on these values and the aforementioned prevalence estimates for diseased lesions, we can calculate the PPV and NPV (Section 4.3.3) for our test: 55.0% and 95.0%, respectively. These results indicate that the spectral diagnostic algorithms would be useful in distinguishing lesions that do not require biopsy (or further evaluation) from those that do. Therefore, spectroscopy, may serve to reduce the number of unnecessary biopsies for patients presenting with suspicious lesions. Spectroscopy could be used as a rapid screening tool to indicate whether a person being seen by a dentist needs further evaluation.

6.4 Conclusions

The goal of this work was to use physical parameters derived from reflectance and fluorescence spectra to create spectral diagnostic algorithms for two types of applications. The first algorithms were designed to distinguish visible lesions from visibly healthy mucosa. The second set was designed to distinguish benign lesions from dysplastic/malignant lesions. The data and findings from the HV study were incorporated into the analysis and used to address the spectral variations due to anatomy. Based on the spectral parameters we extracted from healthy volunteers, the following groups of sites were identified as the most spectrally similar: 1) BM and SP, 2) FM and VT, 3) DT and RT, and GI and HP.

Distinctions between clinically normal and clinically abnormal tissue areas can easily be distinguished based on spectral parameters, particularly for algorithms developed from and applied to individual sites or spectrally similar sites. In addition, the distinction between healthy mucosa and visible lesions is independent of the presence of disease. This was shown by the fact that the results for healthy mucosa versus benign lesions were comparable to those for healthy mucosa versus dysplastic/malignant lesions.

The more clinically relevant separation of determining whether a clinically abnormal site is benign or dysplastic/malignant, was considerably more challenging. Just as for the first type of comparison, diagnostic algorithms for distinguishing benign lesions from dysplastic/malignant lesions were generally more successful when developed for a specific anatomic site or group of spectrally similar sites.

We have demonstrated that the spectral differences among anatomic sites can impact the performance of both types of diagnostic algorithms and presented a strategy for overcoming this limitation based on the concept of spectrally similar sites. Furthermore, we demonstrate that comparisons between healthy mucosa and lesions are not a reliable means of assessing the potential of spectroscopy to distinguish benign from dysplastic/malignant lesions because the diagnostic parameters are almost completely different for the two types of comparisons. These results emphasize the need to treat clinically abnormal and clinically normal mucosa as separate entities and to develop separate diagnostic algorithms for each case. Our results indicate the spectral diagnostic algorithms which we have developed and applied to specific sites can reduce the number of unnecessary biopsies, and thus serve as a screening tool, and test for ruling out the need for biopsy when a patient presents with a suspicious lesion in the clinic.

Future work should focus on high risk sites for oral cancer, such as the tongue (especially LT and VT), floor of the mouth, and soft palate. Since lesions are most likely to be encountered at these sites, successful diagnostic algorithms developed specifically for these sites may provide a significant clinical benefit.

6.6 References

1. J. Pindborg, P. Reichart, C. Smith and I. van der Waal, *World Health Organization: histological typing of cancer and precancer of the oral mucosa*, Springer-Verlag, Berlin, 1997
2. B. Rosner, *Fundamentals of biostatistics*, Duxbury, Pacific Grove, CA, 2000
3. A. K. Jain, R. P. W. Duin and J. C. Mao, "Statistical pattern recognition: A review," *Ieee Transactions on Pattern Analysis and Machine Intelligence*, **22**(1), 4-37 (2000).
4. S. Silverman, "Demographics and occurrence of oral and pharyngeal cancers - The outcomes, the trends, the challenge," *Journal of the American Dental Association*, **132**(7S-11S (2001).
5. D. L. Heintzelman, U. Utzinger, H. Fuchs, A. Zuluaga, K. Gossage, A. M. Gillenwater, R. Jacob, B. Kemp and R. R. Richards-Kortum, "Optimal excitation wavelengths for in vivo detection of oral neoplasia using fluorescence spectroscopy," *Photochemistry and Photobiology*, **72**(1), 103-113 (2000).
6. S. K. Majumder, S. K. Mohanty, N. Ghosh, P. K. Gupta, D. K. Jain and F. Khan, "A pilot study on the use of autofluorescence spectroscopy for diagnosis of the cancer of human oral cavity," *Current Science*, **79**(8), 1089-1094 (2000).
7. S. K. Majumder, N. Ghosh, S. Kataria and P. K. Gupta, "Nonlinear pattern recognition for laser-induced fluorescence diagnosis of cancer," *Lasers in Surgery and Medicine*, **33**(1), 48-56 (2003).
8. S. P. Schantz, V. Kolli, H. E. Savage, G. P. Yu, J. P. Shah, D. E. Harris, A. Katz, R. R. Alfano and A. G. Huvos, "In vivo native cellular fluorescence and histological characteristics of head and neck cancer," *Clinical Cancer Research*, **4**(5), 1177-1182 (1998).
9. C. A. Waldron and W. G. Shafer, "Leukoplakia Revisited - Clinicopathologic Study 3256 Oral Leukoplakias," *Cancer*, **36**(4), 1386-1392 (1975).
10. M. G. Muller, T. A. Valdez, I. Georgakoudi, V. Backman, C. Fuentes, S. Kabani, N. Laver, Z. M. Wang, C. W. Boone, R. R. Dasari, S. M. Shapshay and M. S. Feld, "Spectroscopic detection and evaluation of morphologic and biochemical changes in early human oral carcinoma," *Cancer*, **97**(7), 1681-1692 (2003).
11. A. Gillenwater, R. Jacob, R. Ganeshappa, B. Kemp, A. K. El-Naggar, J. L. Palmer, G. Clayman, M. F. Mitchell and R. Richards-Kortum, "Noninvasive diagnosis of oral neoplasia based on fluorescence spectroscopy and native tissue autofluorescence," *Archives of Otolaryngology-Head & Neck Surgery*, **124**(11), 1251-1258 (1998).
12. H. J. van Staveren, R. L. P. van Veen, O. C. Speelman, M. J. H. Witjes, W. M. Star and J. L. N. Roodenburg, "Classification of clinical autofluorescence spectra of oral leukoplakia using an artificial neural network: a pilot study," *Oral Oncology*, **36**(3), 286-293 (2000).
13. A. Amelink, O. P. Kaspers, H. Sterenborg, J. E. van der Wal, J. L. N. Roodenburg and M. J. H. Witjes, "Non-invasive measurement of the

morphology and physiology of oral mucosa by use of optical spectroscopy," *Oral Oncology*, **44**(1), 65-71 (2008).

14. K. A. Steinmetz and J. D. Potter, "Vegetables, fruit, and cancer prevention: A review," *Journal of the American Dietetic Association*, **96**(10), 1027-1039 (1996).
15. S. T. Mayne, B. Cartmel, M. Baum, G. Shor-Posner, B. G. Fallon, K. Briskin, J. Bean, T. Z. Zheng, D. Cooper, C. Friedman and W. J. Goodwin, "Randomized trial of supplemental beta-carotene to prevent second head and neck cancer," *Cancer Research*, **61**(4), 1457-1463 (2001).
16. J. C. Rhee, F. R. Khuri and D. M. Shin, "Advances in chemoprevention of head and neck cancer," *Oncologist*, **9**(3), 302-311 (2004).
17. A. M. Y. Nomura, R. G. Ziegler, G. N. Stemmermann, P. H. Chyou and N. E. Craft, "Serum micronutrients and upper aerodigestive tract cancer," *Cancer Epidemiology Biomarkers & Prevention*, **6**(6), 407-412 (1997).
18. D. C. G. de Veld, M. Skurichina, M. J. H. Wities, R. P. W. Duin, H. J. C. M. Sterenborg and J. L. N. Roodenburg, "Autofluorescence and diffuse reflectance spectroscopy for oral oncology," *Lasers in Surgery and Medicine*, **36**(5), 356-364 (2005).
19. C. Y. Wang, T. Tsai, H. M. Chen, C. T. Chen and C. P. Chiang, "PLS-ANN based classification model for oral submucous fibrosis and oral carcinogenesis," *Lasers in Surgery and Medicine*, **32**(4), 318-326 (2003).
20. R. J. Mallia, S. S. Thomas, A. Mathews, R. R. Kumar, P. Sebastian, J. Madhavan and N. Subhash, "Laser-induced autofluorescence spectral ratio reference standard for early discrimination of oral cancer," *Cancer*, (2008).
21. *Guide to clinical preventive services*", Baltimore, 1996, www.ahrq.gov/clinic/cpsix.htm May 21, 2008
22. L. Dobrossy, "Epidemiology of head and neck cancer: Magnitude of the problem," *Cancer and Metastasis Reviews*, **24**(1), 9-17 (2005).
23. C. Scheifele, P. A. Reichart and T. Dietrich, "Low prevalence of oral leukoplakia in a representative sample of the US population," *Oral Oncol*, **39**(6), 619-25 (2003).

CHAPTER 7

Discussion and Conclusions

In this chapter, we begin by discussing our progress towards the initial goals, and conclude with a statement on the major achievements of this work.

7.1 Accomplishments of this Thesis

The goal of this project was to use reflectance and fluorescence spectroscopy to non-invasively and quantitatively identify pre-cancerous or cancerous changes in the oral cavity. In Chapter 1 we presented the four objectives for the investigation, which were as follows:

1) To advance the instrumentation for spectral data collection
2) To develop spectral models for the accurate modeling of reflectance and fluorescence spectra from oral tissue and the reliable extraction of parameters relating to tissue morphology and biochemistry
3) To collect *in vivo* reflectance and fluorescence spectra from healthy oral tissue in healthy volunteers, and relate the extracted physical parameters to tissue anatomy
4) To collect *in vivo* reflectance and fluorescence spectra from patients with lesions and develop spectral diagnostic algorithms for identifying and categorizing oral lesions

Here we discuss our accomplishments and progress towards these objectives.

7.1.1 *Advances in Instrumentation*

In Chapter 4 we described the advances in instrumentation made as a result of the elucidation of the source of an artifact in the measured reflectance. Through a series of experiments with phantoms we identified second order light as the cause of an increase in the reflectance measured by the instrument at longer wavelengths (>~550 nm) as a result of a change to grade UV optics. A method to correct the artifact was then developed which permitted data previously collected to be recover. Furthermore, we were able to implement a change in the FastEEM hardware to ensure accurate detection of future data, by collecting an unfiltered and filtered reflectance spectrum during each set of measurements.

7.1.2 *Spectral Model Development*

There were several additions to the basic models developed by our laboratory to model the reflectance and fluorescence spectra. One of the critical improvements in the application of the reflectance model, was to take vessel packaging into account when fitting the hemoglobin absorption features. Before this addition, significant deviations between the data and modeled fit were frequently noted. This addition also provided the extraction of a new reflectance parameter, the effective vessel radius.

Before this work, optical spectroscopic studies performed by our laboratory and by other investigators had not recognized β-carotene as a significant absorber in oral tissue. The discovery of this new absorber was the second major improvement in our modeling of the reflectance data. This finding was made in large part because of

the large data sets collected and analyzed in the present work. Whereas small deviations in the quality of the fit that appear in smaller data sets may not be fully appreciated, the systematic nature of these differences can become evident over the course of examining hundreds of spectra. In a similar manner, we discovered that the fluorescence spectra collected at 340 nm excitation was best modeled as a combination of three fluorophores rather than two. This permitted accurate modeling of the fluorescence emission at this excitation for all the samples. We attribute the two absorbers to two different types of collagens.

7.1.3 In Vivo Collection of Spectral Data from Healthy Volunteers

Oral tissue exhibits a range of architectural characteristics among the various anatomic sites. One notable example is that while most of the oral cavity is non-keratinized, the GI, HP and some regions of the DT are keratinized. Other key differences include the thickness of the epithelium (i.e. BM>>FM), the composition of the submucosa (i.e. bone, muscle), the density and size of the vessels, as well as the presence of other unique features such as glands (i.e. BM,HP,SP) and papillae (DT). We expect these variations to also be reflected in the spectral information we collect from the tissue.

In the healthy volunteer study we used reflectance and fluorescence spectroscopy to identify and characterize 9 different anatomic sites in the oral cavity. The morphological and biochemical information provided by the extracted spectral parameters were then related to the physical properties of the tissue. For the majority of the spectral parameters, certain sites or groups of sites were observed to exhibit

distinct values from one another. K-means clustering analysis was performed to identify sites or groups of sites which showed similar features and revealed that sites which are normally keratinized, most notably the hard palate and gingiva, were distinct from non-keratinized sites for most parameters and clustered together. We also tested whether the 95% confidence interval for the distribution of each parameter, extracted from a subset of the normal tissue data, could correctly characterize a second validation data set. Excellent classification accuracy was demonstrated. Our results revealed that intrinsic differences among sites exist. Furthermore, differences in anatomy are reflected in the extracted spectral parameters, and can produce separations among sites. The results of the study provided strong evidence that a robust and accurate spectroscopic-based diagnostic algorithm for oral cancer will need to be applied in a site specific manner to improve the diagnostic capability, and to ensure accurate evaluation of malignancy while avoid the confounding effects of anatomic properties.

7.1.4 In Vivo Collection of Spectral Data from Patients

The goal of the patient study was to develop spectroscopic diagnostic algorithms with which we can non-invasively distinguish healthy mucosa from lesions (clinically abnormal mucosa), and benign lesions from dysplastic/malignant lesions. Furthermore, we tested the concept of developing and applying the diagnostic algorithms specifically for spectrally similar sites, and to compare the performance. We compared the performance of the algorithms using this approach to that obtained when all sites were combined, or only individual sites were examined.

Our findings demonstrated clear spectral differences between clinically normal tissue and lesions. For this comparison, the ROC-AUC values were highest (>0.90) for diagnostic algorithms developed and applied either to a single anatomic site or to spectrally similar sites. Spectral diagnostic algorithms developed to distinguish between benign lesions and dysplastic/malignant lesions were not as successful, as for comparisons between healthy mucosa and lesions. ROC-AUC values between 0.60 and 0.75 were obtained. The most successful algorithms were developed for the LT and the combination of the FM and VT. Furthermore, the diagnostic parameters for this comparison were substantially different from those used to distinguish healthy mucosa from lesions.

Once again, anatomy had a strong impact on the results of the analysis. We demonstrated that the spectral differences among anatomic sites can impact the performance of both types of diagnostic algorithms. From the data we conclude that the approach of developing and applying spectral diagnostic algorithms to individual or spectrally similar sites improves diagnostic performance, while overcoming the challenge of limited sample numbers.

7.2 Conclusions

In this thesis we have presented our investigation on the use of reflectance and fluorescence spectroscopy for the detection of oral dysplasia/malignancy. Our goal was to use these non-invasive techniques to evaluate both the structural and

biochemical properties of the tissue and to identify a spectroscopic "signature" for dysplasia/malignancy.

In this investigation we have taken preliminary steps towards this goal by examining the impact of tissue anatomy and exploring the optimal means of diagnostic algorithm development. We have investigated the impact of tissue anatomy for 9 sites in the oral cavity and have fully characterized the spectral profile of these sites. We can successfully distinguish healthy mucosa from clinically abnormal mucosa using the spectral parameters. Finally, we have successfully created diagnostic algorithms to distinguish benign lesions from dsyplastic/malignant lesions for the LT, and the FM/VT.

This work has highlighted the complexity and challenges of using spectroscopic approaches to identify pathology in the oral cavity as a result of the innate heterogeneity in the anatomy of this site and we have suggested some approaches for how to realize a successful diagnostic technology. Nevertheless, the findings we have presented show significant progress towards this goal, and serve as an important guide for how to realize a successful spectral diagnostic technology.

CHAPTER 8

Future Directions

In this final chapter of the thesis, we propose some future directions for subsequent investigations on the spectroscopic detection of oral cancer.

8.1 Future Directions

8.1.1 Clinical Studies

Based on the results presented in this thesis, the next step in this project would be to test the diagnostic algorithms developed for specific sites in Chapter 6 prospectively in a large scale clinical study. We could test the ability of spectroscopy to indicate whether a biopsy is not needed (benign lesion), when a patient presents with a suspicious lesion, thus reducing the number of unnecessary biopsies.

In addition to evaluating visible lesions, the margins of excised lesions could also be assessed for invisible disease, by comparing the parameters extracted to those obtained for healthy mucosa (HV data). Another database of spectral data and correlated biopsies collected from clinically normal mucosa with benign, dysplastic, or malignant changes according to pathology would be needed to create spectral diagnostic algorithms to distinguish benign from dysplastic/malignant lesions. Evaluating the margins of excised lesions for invisible disease can also be performed efficiently for a large area of tissue using a wide-scale imaging device, as described in Section 8.1.3.

We may further improve upon the work described here by including additional excitation wavelengths, developing a normalization procedure to reduce inter-patient variability, such as normalizing data by the contralateral site in the patient, and developing the software to enable real-time diagnosis.

8.1.2 Probe Design

The diffuse reflectance and fluorescence spectra measured by the FastEEM reflect not only the tissue architecture and biochemistry, but a number of other factors. In particular, since the excitation light is delivered and collected via an optical fiber, the measured signal is impacted by the geometry of the delivery and collection fibers[1,2], including the numerical aperture (NA), diameter of the fibers, spacing between the delivery and collection fibers, and collection efficiency[3]. A number of probe designs have been explored in order to optimize the collection of light returning from a specific depth within the tissue or differentiate the signal returned from various layers[4-9]. Common approaches include varying the separation between the source and detector fibers and using an angled delivery/collection fiber. One can then isolate light emitted from the structures/regions of interest, or from layers (e.g. epithelial layer, stroma) in which the contrast in the spectral signal between normal and malignant tissue is greatest, leading to a enhanced sensitivity to disease-related changes. The work presented in this thesis was carried out using a single probe design. It is possible that with some of the modifications described above, we may be able to enhance our sensitivity to disease, and even filter out the effect of keratin.

8.1.3 Quantitative Spectroscopy Imaging (QSI) Instrument

Despite the detailed information and relative technical simplicity of the point probe technology, an imaging device offers several advantages for the detection of early disease: 1) A wide area of tissue can be inspected including both lesions and

margins of excised lesions 2) It would facilitate the clinician's ability to monitor a particular site for disease development with time, thus avoiding the need for unnecessary biopsies 3) Variability introduced into the measured signal due to the effect of probe pressure is no longer an issue.

A non-contact wide-scale imaging device, known as the quantitative spectroscopy imaging (QSI) instrument, has been designed for early cancer detection in the oral cavity and cervix. The instrument was designed based on the concept of a "virtual probe", in which 1 mm spots of the excitation light are raster scanned along an approximately 2 cm by 2 cm area of tissue. Figure 8.1 shows examples of parameter maps for 1 cm by 1cm area of tissue obtained after applying the spectral

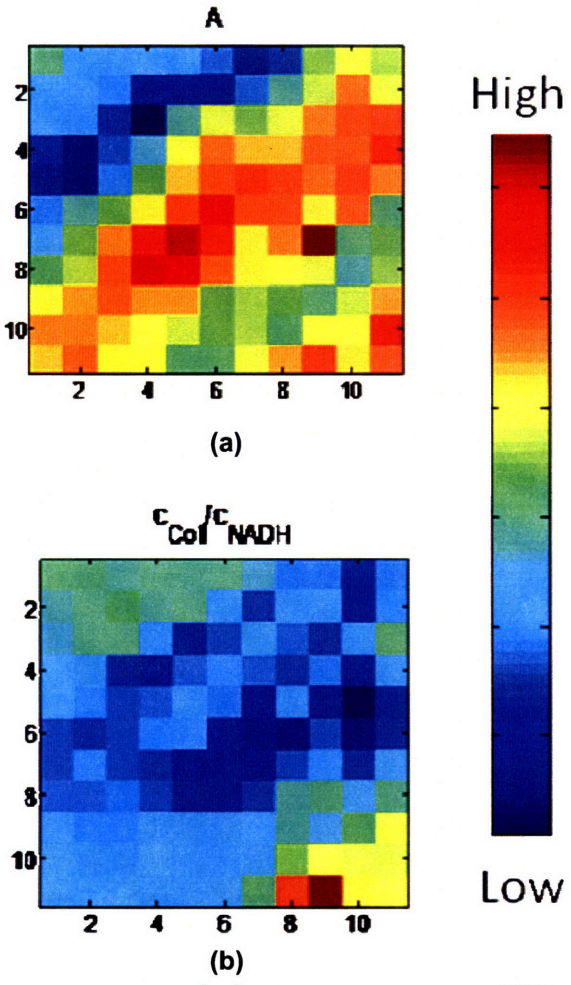

Figure 8.1 Examples of parameter maps generated from modeling data collected by the QSI instrument from an oral lesion on the ventral surface of the tongue. The units are shown in millimeters (a) A map of the A parameter (b) A map of the ratio of collagen to NADH (reduced nicotinamide adenine dinucleotide)

models to *in vivo* image data collected from an area of leukoplakia on the ventral surface of the tongue. Each spot is analogous to a single measurement collected by an optical fiber probe, therefore the knowledge and experience gained from the FastEEM clinical instrument and the DRS and IFS modeling strategies can be applied to the analysis of the measured data.

The first generation instrument is currently being re-designed as an endoscopic system based on an imaging fiber bundle. The fiber bundle could be advanced through the accessory channel of an endoscope and additional sites along the upper aerodigestive tract, including the pharynx and larynx, are accessible to examination.

8.2 References

1. P. R. Bargo, S. A. Prahl and S. L. Jacques, "Optical properties effects upon the collection efficiency of optical fibers in different probe configurations," **9**(2), 314-321 (2003).
2. P. R. Bargo, S. A. Prahl and S. L. Jacques, "Collection efficiency of a single optical fiber in turbid media," **42**(16), 3187-3197 (2003).
3. P. R. Bargo, S. A. Prahl, T. T. Goodell, R. A. Sleven, G. Koval, G. Blair and S. L. Jacques, "In vivo determination of optical properties of normal and tumor tissue with white light reflectance and an empirical light transport model during endoscopy," **10**(3), (2005).
4. A. Amelink, M. P. L. Bard, S. A. Burgers and H. J. C. M. Sterenborg, "Single-scattering spectroscopy for the endoscopic analysis of particle size in superficial layers of turbid media," *Applied Optics*, **42**(19), 4095-4101 (2003).
5. A. Wang, V. Nammalavar and R. Drezek, "Experimental evaluation of angularly variable fiber geometry for targeting depth-resolved reflectance from layered epithelial tissue phantoms," *Journal of Biomedical Optics*, **12**(4), - (2007).
6. D. Arifler, R. A. Schwarz, S. K. Chang and R. Richards-Kortum, "Reflectance spectroscopy for diagnosis of epithelial precancer: model-based analysis of fiber-optic probe designs to resolve spectral information from epithelium and stroma," *Applied Optics*, **44**(20), 4291-4305 (2005).
7. D. Hattery, B. Hattery, V. Chernomordik, P. Smith, M. Loew, J. Mulshine and A. Gandjbakhche, "Differential oblique angle spectroscopy of the oral epithelium," **9**(5), 951-960 (2004).
8. L. Nieman, A. Myakov, J. Aaron and K. Sokolov, "Optical sectioning using a fiber probe with an angled illumination-collection geometry: evaluation in engineered tissue phantoms," *Applied Optics*, **43**(6), 1308-1319 (2004).
9. M. C. Skala, G. M. Palmer, C. F. Zhu, Q. Liu, K. M. Vrotsos, C. L. Marshek-Stone, A. Gendron-Fitzpatrick and N. Ramanujam, "Investigation of fiber-optic probe designs for optical spectroscopic diagnosis of epithelial pre-cancers," *Lasers in Surgery and Medicine*, **34**(1), 25-38 (2004).

List of Abbreviations

α	Oxygen saturation
ε	Extinction coefficient
AUC	Area under the curve
BM	Buccal mucosa
BT	Base of the tongue
$c\beta C$	Concentration of β-carotene
cHb	Concentration of hemoglobin
C.I.	Confidence interval
CIS	Cancer *in situ*
CRM	Confocal reflectance microscopy
DRS	Diffuse reflectance spectroscopy
DT	Dorsal surface of the tongue
FAD	Flavin adenine dinucleotide
FastEEM	Fast excitation emission matrix
FM	Floor of the mouth
FWHM	Full width at half-maximum
GI	Gingiva
Hb	Deoxyhemoglobin
HbO_2	Oxyhemoglobin
H&E	Hemotoxylin and eosin
HP	Hard palate
HV	Healthy volunteer
IFS	Intrinsic fluorescence spectroscopy
IL	Lips
IQR	Interquartile range
LCV	Leave-one-out cross-validation
LLR	Log-likelihood ratio test
LSS	Light scattering spectroscopy
LT	Lateral surface of the tongue
LX	Larynx
MCR	Multivariate curve resolution
MLE	Maximum likelihood estimation
MPM	Multi-photon microscopy
NA	Numerical aperture
NADH	Nicotinamide adenine dinucleotide (reduced)
NPV	Negative predictive value
PCA	Principal component analysis
PPIX	Protoporphyrin IX
PPV	Positive predictive value
ROC	Receiver operator characteristic
RSD	Relative standard deviation
RT	Retromolar trigone
SCC	Squamous cell carcinoma
SNR	Signal-to-noise ratio
SP	Soft palate
TL	Tonsils
TMS	Tri-modal spectroscopy
VT	Ventral surface of the tongue